



PS Claim 8; Fig 1A-C; 27pp; English.

XX This sequence represents a novel human glutathione S-transferase, GSTS.

CC Nucleotide sequences encoding GSTS were initially isolated from a urologic cDNA library, and subsequently extended using cDNA libraries derived from other tissues, such as brain or bladder. The present sequence is encoded by a consensus cDNA. GSTS, and nucleotides which encode it may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate GSTS expression, such as cancers and immune disorders. Nucleotides which encode GSTS may be used in gene therapy to treat disorders associated with reduced expression or activity of GSTS, and in antisense therapy for disorders associated with increased GSTS expression or activity. They may also be used for the recombinant production of GSTS, and as a source of probes and primers to detect and quantitate the presence of similar nucleic acid sequences, particularly for the diagnosis of GSTS-associated disorders. GSTS proteins may be used as antigens in the production of antibodies against GSTS and in assays to identify modulators (agonists and antagonists) of GSTS expression and activity. The anti-GSTS antibodies and GSTS antagonists may also be used to downregulate GSTS expression and activity. Antagonists of GSTS expression and function may be used to treat immune disorders (e.g. AIDS, anaemia, asthma, Crohn's disease, irritable bowel syndrome, multiple sclerosis, osteoarthritis, microbial infections and cancers (e.g., leukaemia, lymphoma, melanoma, and cancers of the breast, prostate, liver, lung and brain). The anti-GSTS antibodies may also be used as diagnostic agents.

XX Sequence 226 AA;

Query Match 100.0%; Score 226; DB 21; Length 226;  
Best Local Similarity 100.0%; Pred. No. 2.5e-219;  
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGPLPRTVELFDYDVLSPYSPWLGFETLCRYONIWNINQLRPLSLITGIMKDSGNKPPGLP 60  
Db 1 mgplprtvelfdydvspyswlgfetlcryoniniqlrplslitgimkdsgnkppglp 60  
Qy 61 RKGLYMANDLKLRRHLHQIPTHPKPKFLSVMLFEKGSLSAMRFLTAVNLEHPEMLEKASRE 120  
Db 61 rkglymandlklrrhlhqiphtpkflsvmlfeekgslsamrfltavnlehmekasre 120  
Qy 121 LWMRWWSRNEDITEPOSILAAEAKAGMSAEQAGLLEKIAATPKVKNQKETTEAACRYGA 180  
Db 121 lwmrwssrneditpeosilaaeakagmsaeqagllekiaatpkvknqketteaacryga 180  
Qy 181 FGLPITVAHDGQPHMLFGSDRMELLAHLIGEKWMGPPIPNAVRL 226  
Db 181 fglpitvahdgqphmlfgsdrmellahligeckwmgpippavnrl 226

RESULT 2  
Y59988 standard; Protein: 256 AA.  
XX DE19811948-A1.  
XX KW Human; tumour; cancer; anticancer; cytostatic; EST:  
KW treatment; uterine; gene therapy; expressed sequence tag.  
XX OS Homo sapiens.  
XX PN DE19811948-A1.  
XX PD 21-OCT-1999.  
XX PR 17-APR-1998; 98DE-1017948.  
PR 17-APR-1998; 98DE-1017948.

XX PA (META-) METAGEN GES GENOMFORSCHUNG MBH.  
XX PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;  
XX DR WPI: 1999-591957/51.  
XX DR N-FSDB; Z41996.  
XX PS New nucleic acid sequences expressed in uterine cancer tissues, and derived polypeptides, for treatment of uterine and endometrial cancer  
XX PT and identification of therapeutic agents.  
XX PS Claim 23; Page 294, 444PB; German.  
XX This invention describes novel human nucleic acid (cDNA) sequences (A), that are highly expressed in uterine tumour tissue and which have anticancer and cytostatic activity. (A) are used (i) for recombinant expression of polypeptides (B) and (ii) to isolate complete genes. (B) are used (i) to identify agents suitable for treatment of uterine or endometrial cancer; (ii) directly for treating these forms of cancer (including expression from gene therapy vectors) and (iii) for generation of specific antibodies. (A) are identified by assembling ESTs (expressed sequence tags) from a particular tissue type before comparison of expression patterns. This allows a significantly longer fragment of the gene to be revealed, so should reduce the number of failures associated with the fact that ESTs from different libraries may represent different parts of the same unknown gene, distorting the estimated frequency of occurrence in a particular tissue. Y59911-Y60328 represent protein EST fragments encoded by the human endometrium tumour cDNA library derived from the human endometrium tumour cDNA library derived in Z41981-242121.  
XX SQ Sequence 256 AA;  
XX Query Match 100.0%; Score 226; DB 20; Length 256;  
Best Local Similarity 100.0%; Pred. No. 2.8e-219;  
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MGPLPRTVELFDYDVLSPYSPWLGFETLCRYONIWNINQLRPLSLITGIMKDSGNKPPGLP 60  
Db 31 mgplprtvelfdydvspyswlgfetlcryoniniqlrplslitgimkdsgnkppglp 90  
Qy 61 RKGLYMANDLKLRRHLHQIPTHPKPKFLSVMLFEKGSLSAMRFLTAVNLEHPEMLEKASRE 120  
Db 91 rkglymandlklrrhlhqiphtpkflsvmlfeekgslsamrfltavnlehmekasre 150  
Qy 121 LWMRWWSRNEDITEPOSILAAEAKAGMSAEQAGLLEKIAATPKVKNQKETTEAACRYGA 180  
Db 151 lwmrwssrneditpeosilaaeakagmsaeqagllekiaatpkvknqketteaacryga 210  
Qy 181 FGLPITVAHDGQPHMLFGSDRMELLAHLIGEKWMGPPIPNAVRL 226  
Db 211 fglpitvahdgqphmlfgsdrmellahligeckwmgpippavnrl 226  
RESULT 3  
G01523 standard; Protein: 131 AA.  
XX AC G01523;  
XX DT 06-OCT-2000 (first entry)  
XX DE Human secreted protein, SEQ ID NO: 5604.  
XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping.  
XX OS Homo sapiens.  
XX PN EP1033401-A2.  
XX PD 06-SEP-2000.

XX	21-FEB-2000;	2000EP-0200610.	PD	28-SEP-2000.
PF			XX	
XX	26-FEB-1999;	99US-0122487.	PF	16-MAR-2000;
PR			XX	2000060-US06824.
XX			PR	19-MAR-1999;
(GEST ) GENSET.			PR	99US-0125359.
PA			XX	03-DEC-1999;
XX			XX	99US-0168664.
Dumas Milne Edwards J;	Ducleart A;	Giordano J;	PA	(HUMA-) HUMAN GENOME SCI INC.
PI			XX	
XX			Rosen CA;	Ruben SM;
WPI; 2000-500381/45.			Komatsoulis G;	
DR .			PI	
N-PSDB; C01529.			XX	
XX			WPL; 2000-594574/56.	
PT			DR	N-PSDB; C60069.
PT	New nucleic acid that is a 5' expressed sequence tag (5' EST) for obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for diagnostic, forensic, gene therapy and chromosome mapping procedures -	XX	Human secreted proteins and gene sequences encoding them, useful for detection prevention, and treatment of various disorders such as cancer and immune system disorders -	
PT			PT	
XX			XX	
PS	Claim 13: SEQ ID 5604; 71pp + CD-ROM; English.	XX	Disclosure; Page 83; 442pp; English.	
XX			PS	
CC	The present sequence is a polypeptide encoded by one of a large number of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were prepared from total human RNAs or poly+ RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors.	XX	The polynucleotide sequences given in C60025-C60071 encode the human secreted proteins represented in B3484-B34900. Sequences B34901-B34976 are fragments of proteins encoded by the genes, and also proteins with which they share sequence homology. The proteins have activities based on the tissues in which their encoding genes are expressed. Examples of the proteins activities include: neuroprotective; cytostatic; cardioactive; immunomodulatory; general muscular activity; pulmonary; nephrotropic; antiinfective; gynaecological; and antibacterial. The human secreted proteins, polynucleotides, antagonists and antagonists of the invention may be useful in treating, preventing and/or diagnosing the various diseases, disorders and conditions such as neural, immune, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal and proliferative disorders and cancer. They may also be used in the treatment of wounds, and infectious diseases. The polypeptides may be used as a food additive or preservative to increase storage capabilities. Sequences C60016-C60024 and B3483 are used in the course of the invention during the identification and characterisation of the protein and nucleotide sequences.	
SQ	Sequence 131 AA;	XX	Sequence 97 AA;	
Query Match	58.0%	Score 131;	DB 21;	Length 97;
Best Local Similarity	100.0%	Pred. No. 5.4e-124;		
Matches	131;	Conservative 0;	Mismatches 0;	Gaps 0;
Db	1 mgplprtvelfdydlyspwysmglfeliyclqnyiwnnlnqlreslitzgimkdsqnkppglp 60	Qy	1 MGPLPRTVLFYDLYSPWYSMGLFELIYCLQNYIWNNLNQLRESLITZGIMKDSQNKPPLGP 60	Query Match 42.9%; Score 97; DB 21;
Db	1 mgplprtvelfdydlyspwysmglfeliyclqnyiwnnlnqlreslitzgimkdsqnkppglp 60	Db	1 mgplprtvelfdydlyspwysmglfeliyclqnyiwnnlnqlreslitzgimkdsqnkppglp 60	Local Similarity 100.0%; Pred. No. 6.5e-90; Mismatches 0; Indels 0; Gaps 0;
Qy	61 rkgllymandjklrrhlqlipthpkdflyspwysmglfeliyclqnyiwnnlnqlreslitzgimkdsqnkppglp 120	Qy	61 rkgllymandjklrrhlqlipthpkdflyspwysmglfeliyclqnyiwnnlnqlreslitzgimkdsqnkppglp 120	Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	61 rkgllymandjklrrhlqlipthpkdflyspwysmglfeliyclqnyiwnnlnqlreslitzgimkdsqnkppglp 120	Db	61 rkgllymandjklrrhlqlipthpkdflyspwysmglfeliyclqnyiwnnlnqlreslitzgimkdsqnkppglp 120	Qy 130 EDITEPOSTILAAEAKAGMSAEQAOGGLEKIAATPKYKNQLKETTEACRYGAGFLPITYWAH 189
Qy	121 lwmrvwsrned 131	Qy	121 lwmrvwsrned 131	Db 1 editepqslaaekagmsaeqaogglekiaatpkknqketeaacrygagflpitywa 60
Db	121 lwmrvwsrned 131	Db	121 lwmrvwsrned 131	Qy 190 VDGOTHMILEGSDRMELLAHLLGEKWMGPPIPPAVNRL 226
RESULT	4		Db 61 vdqgthmlfgsdrmellahllgekwmgpipavaarl 97	Db 61 vdqgthmlfgsdrmellahllgekwmgpipavaarl 97
B34972		RESULT 5		
ID	B34972 standard; Protein: 97 AA.	Id	B43298 standard; Protein: 300 AA.	
XX		XX		
AC	B34972;	XX		
DT	26-JAN-2001 (first entry)	XX		
XX				
DE	Human secreted protein sequence encoded by gene 45 SEQ ID NO:176.	XX		
XX				
Human; secreted protein; neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular; vulnerability; nephrotropic; antineoplastic; gynaecological; antibacterial; antiarthritic; immunosuppressive; cardiotonic; anticonvulsant; osteopathic; coagulant; vasotonic; immunostimulant; thrombolytic; antiinflammatory; hypotensive; dermatological; antihematic; antifungal; antidiabetic; antiviral; antibacterial; antihematic; antithyroid; antidiabetic; antianemic; gene therapy; cancer; proliferative disorder; hypertension;	XX			
KW				
OS	Homo sapiens.	XX		
XX				
PN	WO200056766-A1.	XX		

KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; contraceptive; coagulation;  
 KW thrombosis; contrageptive.  
 XX Homo sapiens.  
 XX WO20058473-A2.  
 XX PD 05-OCT-2000.  
 XX PR 31-MAR-2000; 2000WO-US08621.  
 XX PR 31-MAR-1999; 99US-0127607.  
 XX PR 02-APR-1999; 99US-0127636.  
 XX PR 05-APR-1999; 99US-0127728.  
 XX PR 30-MAR-2000; 2000US-0540763.  
 XX PA (CURA-) CURAGEN CORP.  
 XX PI Shimkets RA, Leach M;  
 XX DR WPI: 2000-602362/57.  
 XX DR N-PSDBB; C7507.  
 XX PT Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease -  
 XX PS Claim 11; Page 5304-5305; 5507pp; English.  
 XX CC C74416 to C77605 encode the proteins given in B40237 to B43397, which  
 CC represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cyostatic; hepatotoxic; pulmonary;  
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;  
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;  
 CC cardiotonic; thrombolytic; coagulant; vasotropin; antidiabetic; hypotensive;  
 CC dermatological; immunosuppressive; ancinflammatory; antibacterial;  
 CC antiviral; antifungal; antiarthematic; antihypertensive; antianæmic. The  
 CC sequences can be used for determining the presence of or predisposition  
 CC to, or preventing or treating pathological conditions associated with an  
 CC ORFX associated disorder. The nucleic acids can be used to express ORFX  
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
 CC used to treat cancers, proliferative disorders, neurodegenerative  
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease, to  
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive.  
 XX Sequence 300 AA;

Query Match 38.9%; Score 88; DB 21; Length 300;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-80;  
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 NEDITEPOSILAAEAKGMSAEQAOGLLEKATPKVKNQLEKETTEACRYGAFLGLPITVA 188  
 Db 213 nediteposilaaeakgmsaeqaoqllekektpkvknqleketteacrygaflglpitva 272

QY 189 HVDGQTAMLFGSDRMELLAHLGERWMG 216  
 Db 273 hvdgqtamlfgsdrmellaahlgerwmg 300

RESULT 6  
 B34971

ID B34971 standard; Protein; 97 AA.  
 XX AC B34971;  
 XX DT 26-JAN-2001 (first entry)  
 DE Gene 45 human secreted protein homologous amino acid sequence #175.  
 XX Human; secreted protein; neuroprotective; cyostatic; cardioactive;  
 KW immunomodulatory; muscular; pulmonary; gastrointestinal; nephrotropic;  
 KW antinefctive; gynaecological; antibacterial; neural disorder; cancer;  
 KW immune disease; reproductive disorder; proliferative disorder;  
 KW gastrointestinal disease; wound healing; infectious disease;  
 KW food additive.  
 XX OS Rattus sp.  
 XX PN WO20056766-A1.  
 XX PR 28-SEP-2000.  
 XX PD 16-MAR-2000; 2000WO-US06824.  
 XX PR 19-MAR-1999; 99US-0125359.  
 XX PR 01-DEC-1999; 99US-0168864.  
 XX PA (HUMA-) HUMAN GENOME SCI INC.  
 XX PI Rosen CA, Ruben SM, Komatsoulis G;  
 XX DR WPI; 2000-594574/56.

Human secreted proteins and gene sequences encoding them, useful for  
 detection, prevention, and treatment of various disorders such as  
 cancer and immune system disorders -  
 Disclosure; Page 83; 442pp; English.

The polynucleotide sequences given in C60025-C60071 encode the human  
 CC secreted proteins represented in B34854-B34900. Sequences B34901-B34976  
 CC are fragments of proteins encoded by the genes, and also proteins with  
 CC which they share sequence homology. The proteins have activities based on  
 CC the tissues in which their encoding genes are expressed. Examples of  
 CC the proteins/activities include: neuroprotective; cyostatic;  
 CC cardioactive; immunomodulatory; general muscular activity; pulmonary;  
 CC general gastrointestinal activity; nephrotropic; antiinfective;  
 CC gynaecological; and antibacterial. The human secreted proteins,  
 CC polynucleotides, antagonists and antagonists of the invention may be  
 CC useful in treating, preventing and/or diagnosing various diseases,  
 CC disorders and conditions such as neural, immune, muscular, reproductive,  
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative,  
 CC disorders and cancer. They may also be used in the treatment of wounds,  
 CC and infectious diseases. The polypeptides may be used as a food additive  
 CC or preservative to increase storage capabilities. Sequences C60016-C60024  
 CC and B34853 are used in the course of the invention during the  
 CC identification and characterisation of the protein and nucleotide  
 CC sequences.

Query Match 5.3%; Score 12; DB 21; Length 97;  
 Best Local Similarity 100.0%; Pred. No. 0.00021;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 MLFGSDRMELLA 207  
 Db 67 mlfgsdrmella 78

RESULT 7  
 B58191

ID B58191 standard; Protein; 179 AA.

XX DT 02-NOV-2000 (first entry)  
 XX DE Human Ma2 protein SEQ ID NO:7.  
 XX DT 14-MAR-2001 (first entry)  
 XX DE Lung cancer associated polypeptide sequence SEQ ID 529.  
 XX KW Human; lung cancer associated protein; neuroprotective; cytostatic;  
 XX KW cardiotropic; immunomodulatory; muscular active; vulnerary;  
 XX KW gastrointestinal; nephrotropic; antiinfective; gynecological;  
 XX KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
 XX KW proliferative disorder; wound healing; infectious disease.  
 OS XX Homo sapiens.  
 XX PN WO200055180-A2.  
 XX PD 21-SEP-2000.  
 XX PR 08-MAR-2000; 2000WO-US05918.  
 XX PF 12-MAR-1999; 99US-0124270.  
 XX PA (HUMA-) HUMAN GENOME SCI INC.  
 XX PA (ROSE/) ROSEN C A.  
 XX PI Ruben SM;  
 XX DR 2000-587514/55.  
 XX DR N-PSDB; F18067.  
 XX PS Lung cancer associated gene sequences, referred to as lung cancer  
 PT antigens, useful for treatment, prevention, and diagnosis of disorders  
 PT such as lung cancer.  
 XX Claim 11; Page 1021; 1425pp; English.  
 XX Polynucleotide sequences F17982 - F18424 encode human lung cancer  
 CC associated proteins represented in B58106 - B58548. Lung cancer  
 CC associated proteins and Polynucleotide sequences, their agonists, and  
 CC antagonists may have neuroprotective; cytostatic; cardioactive;  
 CC immunomodulatory; muscular active general; pulmonary; gastrointestinal  
 CC general; nephrotropic; antiinfective; gynecological; or antibacterial  
 CC activity. The invention also includes antibodies specific for the protein  
 CC or polynucleotide sequences. The lung cancer associated polynucleotide  
 CC sequences may be used for detection of lung cancer, chromosome  
 CC identification, as chromosome markers, and for numerous other diagnostic  
 CC or research purposes. The proteins may be used to treat disorders such as  
 CC neural, immune, muscular, reproductive, gastrointestinal, pulmonary,  
 CC cardiovascular, renal, and proliferative disorders. The proteins may also  
 CC be used in the treatment of wounds and infectious diseases.  
 CC Polynucleotide sequences F18425 - F18433 and peptide B58549 are used in  
 CC the course of the invention for the identification and characterisation  
 CC of the polynucleotide and protein sequences.  
 XX Sequence 179 AA;  
 SQ

Query Match 3 %; Score 8; DB 21; Length 179;  
 Best Local Similarity 100.0%; Pred. No. 3.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 96 SLSAMRFL 103  
 Db 25 s1samlf1 32

RESULT 8  
 B12526 ID B12526 standard; Protein: 195 AA.  
 XX AC B12526;  
 AC B12526 standard; Protein: 195 AA.

Query Match 3.5%; Score 8; DB 21; Length 195;  
 Best Local Similarity 100.0%; Pred. No. 4.1;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 204 ELIAHLIG 211  
 Db 139 eliahlig 146

RESULT 9  
 ID B12528 standard; Protein: 283 AA.  
 XX AC B12528;  
 AC B12528;  
 DT 02-NOV-2000 (first entry)  
 XX DE Human Ma4 protein SEQ ID NO:11.  
 XX KW Ma2; Ma3; Ma4; Ma5; paraneoplastic limbic encephalitis; diagnosis;  
 XX KW paraneoplastic syndrome; neoplasm; colon cancer; brainstem encephalitis;  
 XX KW breast cancer; parotid gland cancer; lung cancer; testicular cancer;  
 XX KW germ-cell tumour.  
 OS XX Homo sapiens.  
 PN JP2000146982-A.

XX	PD	26-MAY-2000.	XX	DR	WPI; 1994-358275/44.
XX	PF	10-NOV-1999;	99JP-0320171.	DR	N-PSDB; Q73061.
XX	XX	10-NOV-1998;	98US-0189527.	XX	Eliciting an immune response to <i>Salmonella</i> - using attenuated <i>Salmonella</i> strains, vector constructs, or compns. Contg.
XX	PR	(SLOK ) SLOAN KETTERING INST CANCER RES.	PT	PT	fimbrial type proteins.
XX	PA	DR	PS	PS	Disclosure; Figure 2A-2D; 95pp; English.
XX	PA	WPI; 2000-468119/41.	XX	XX	The isolated SeFc protein may be used in a vaccine composition to elicit an immune response to <i>Salmonella</i> in animals (e.g. food producing animals) and humans.
XX	DR	DR	CC	CC	
XX	N-PSDB;	A60836.	CC	CC	
XX	XX	Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic encephalitis or neoplasm e.g. colon cancer comprising assessing a test sample for the presence or absence of antibodies to a Ma family polypeptide -	XX	XX	
XX	PT	PS	SQ	Sequence 813 AA;	
XX	PT	Claim 48; Fig 7-8; 27pp; Japanese.	Query Match	3.5%; Score 8; DB 15;	Length 813;
XX	PT	The present invention describes a method for diagnosing a paraneoplastic syndrome or neoplasm. The method comprises assessing a test sample for the presence or absence of antibodies to a Ma family polypeptide (1). The method is used to diagnose a paraneoplastic limbic syndrome especially paraneoplastic limbic encephalitis and/or brainstem encephalitis or neoplasm in an individual. The method diagnoses the encephalitis by assessing antibodies to (1) preferably Ma1, which is indicative of presence of breast cancer, colon cancer, parotid gland cancer, lung cancer, testicular cancer and germ-cell tumours or Ma2, which is indicative of testicular cancer, germ-cell tumour, and lung cancer. The present sequence is the Ma4 protein as given in the present invention.	Best Local Similarity	100.0%; Pred. No. 15;	Length 813;
XX	PT	Sequence 283 AA;	Matches	8; Conservative 0; Mismatches 0;	Indels 0; Gaps 0;
XX	PT	Query Match	QY	166 NOLKETTE 173	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	60 nollette 67	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	11	
XX	PT	59 ellahillg 66	W23573	DT 29-SEP-1997 (first entry)	
XX	PT	Query Match	QY	166 NOLKETTE 173	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	60 nollette 67	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 283 AA;	RESULT	10	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		
XX	PT	Sequence 204 ELLAHILLG 211	RESULT	10	
XX	PT	59 ellahillg 66	R62754	ID R62754 Standard; Protein; 813 AA.	
XX	PT	Query Match	QY	204 ELLAHILLG 211	
XX	PT	3.5%; Score 8; DB 21; Length 283;	Db	59 ellahillg 66	
XX	PT	Best Local Similarity	100.0%; Pred. No. 5.8;		
XX	PT	Matches	8; Conservative 0; Mismatches 0;		



FT sig-peptide 1..22  
 XX Pseudomonas aeruginosa ATCC 33354.  
 OS XX EP357024-A.  
 PN XX PD 07-MAR-1990.  
 XX PF 30-AUG-1989; 89EP-0115992.  
 XX PR 01-SEP-1988; 88DE-3829516.  
 XX PA (BEHW ) BEHRINGWERKE AG.  
 XX PI Duchene M, von Specht U, Domdey H;  
 DR WPI; 1990-068794/10.  
 DR N-PSDB; Q03507.  
 XX Pseudomonas aeruginosa outer membrane lipoprotein - useful for vaccine production.  
 XX PS Example 5; Page 4; 5pp; German.  
 XX CC and its fragments are useful for the production of vaccines and antisera against Pseudomonas infections. The DNA and antibodies are useful for diagnostic purposes.  
 XX SQ Sequence 83 AA;

Query Match 3.1%; Score 7; DB 21; Length 130;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 113 MLEKASR 119  
 Db 76 mlekasr 82

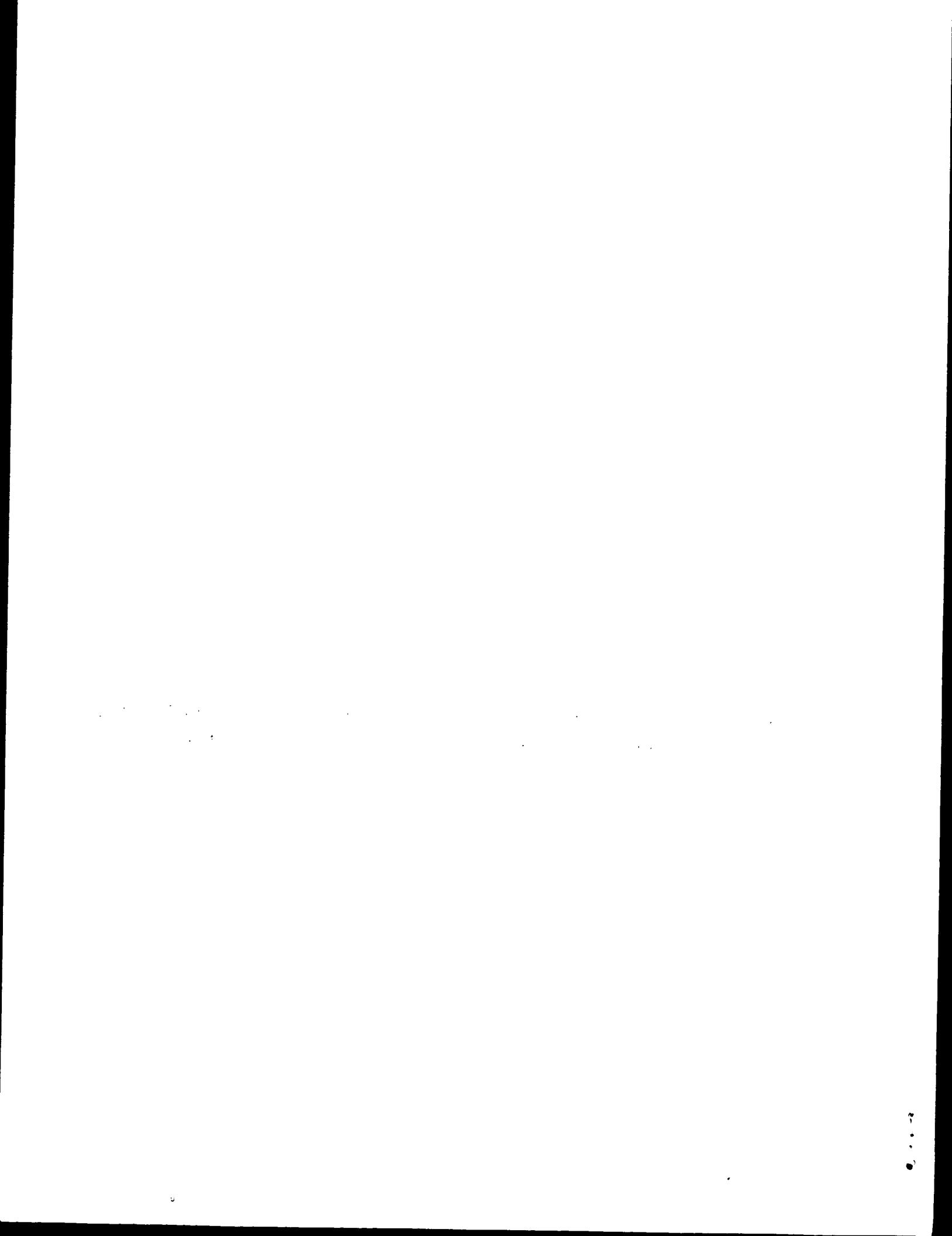
RESULT 15  
 B58516 standard Protein: 130 AA.  
 TD B58516  
 XX AC B58516;  
 XX DT 14-MAR-2001 (first entry)  
 XX DE Lung cancer associated polypeptide sequence SEQ ID 854.  
 KW Human; lung cancer associated protein; neuroprotective; cytostatic;  
 KW cardioactive; immunomodulatory; muscular active; pulmonary;  
 KW gastrointestinal; nephrotropic; antinefertive; gynecological;  
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
 KW proliferative disorder; wound healing; infectious disease.  
 OS Homo sapiens.  
 PN WO200055180-A2.  
 XX PD 21-SEP-2000.  
 XX PF 08-MAR-2000; 2000WO-US05918.  
 XX PR 12-MAR-1999; 99US-0124270.  
 XX PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (ROSE/) ROSEN C A.  
 XX PI Ruben SM;  
 XX DR WPI; 2000-587514/55.  
 DR N-PSDB; F18392.

XX PT Lung cancer associated gene sequences, referred to as lung cancer antigens, useful for treatment, prevention, and diagnosis of disorders such as lung cancer -  
 XX PS Claim 11; Page 1391-1392; 1425pp; English.  
 XX CC Polynucleotide sequences F17982 - F18424 encode human lung cancer associated proteins and polynucleotide sequences in B58106 - B58548. Lung cancer antagonists may have neuroprotective; cytostatic; cardiotropic; immunomodulatory; muscular active; general; vulnerable; gynaecological; or antibacterial activity. The invention also includes antibodies specific for the protein or polynucleotide sequences. The lung cancer associated polynucleotide sequences may be used for detection of lung cancer, chromosome identification, as chromosome markers, and for numerous other diagnostic or research purposes. The proteins may be used to treat disorders such as neural, immune, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal, and proliferative disorders. The proteins may also be used in the treatment of wounds and infectious diseases. The peptides may also be used in the treatment of wounds and infectious diseases.  
 CC Polynucleotide sequences F18425 - F18433 and peptide B58549 are used in the course of the invention for the identification and characterisation of the polynucleotide and protein sequences.  
 XX SQ Sequence 130 AA;

Query Match 3.1%; Score 7; DB 21; Length 130;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 97 LSAMRFL 103  
 Db 1 lsamrfl 7

Search completed: April 19, 2001, 12:58:35  
 Job time: 113 sec





Matches	225;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;	Query Match Score	7.1%;	Score	16;	DB	1;	Length	225;
Qy	2	GPLRTVEFLYDVLSPIVSWIIGFEIILGCRYQNWNINQLRSLITGMDSGNKPGPLPR	61							Best Local Similarity	100.0%;	Pred.	No.	9.9e-09;			
Db	1	GPLRTVEFLYDVLSPIVSWIIGFEIILGCRYQNWNINQLRSLITGMDSGNKPGPLPR	60							Mismatches	0;	Indels	0;				
Qy	62	KGLYMANDKLRLHLQIPTIFPKDPFLSVMLEKGSLSAMRFLTAVNLHPEMLEKASREL	121							Matches	16;	Conservative	0;				
Db	61	KGLYMANDKLRLHLQIPTIFPKDPFLSVMLEKGSLSAMRFLTAVNLHPEMLEKASREL	120							RESULTS	3						
RECA_RAT										RECA_AQUPY							
ID										ID	RECA_AQUPY						
AC										AC	P33542;						
DT										DT	01-FEB-1994	(Rel.	28,	Created)			
DB										DT	01-FEB-1994	(Rel.	28,	Last sequence update)			
Qy	122	WMRWWSRNEDITEPOSLAEEAKAGMSAEQAGGLEKIAITPKVKNOLKETTEAACRYGAF	181							DT	01-OCT-2000	(Rel.	40,	Last annotation update)			
Db	121	WMRWWSRNEDITEPOSLAEEAKAGMSAEQAGGLEKIAITPKVKNOLKETTEAACRYGAF	180							DE	RECA PROTEIN (RECOMBINASE A).						
Qy	182	GLPITVAHVGDQTMHLFGSDRMELLAHLGLGKWMGPIPPVNARL	226							CN	RECA						
Db	181	GLPITVAHVGDQTMHLFGSDRMELLAHLGLGKWMGPIPPVNARL	225							OS	Aquifex pyrophilus.						
										OC	Bacteria; Aquificales; Aquificaceae; Aquifex.						
										OX	NCBI_TaxID=2714;						
										RN	[1]						
										RP	SEQUENCE FROM N.A.						
										RC	STRAIN=KO5A;						
										RX	MEDLINE=95014407;						
										RA	PubMed=7929298;						
										RT	"Cloning, sequencing, and expression of RecA proteins from three distantly related thermophilic eubacteria."						
										RT	J. Biol. Chem. 269:25948-25945 (1994).						
										CC	-!- FUNCTION: CAN CATALYZE THE HYDROLYSIS OF ATP IN THE PRESENCE OF SINGLE-STRANDED DNA, THE ATP-DEPENDENT UPTAKE OF SINGLE-STRANDED DNA BY DUPLEX DNA, AND THE ATP-DEPENDENT HYBRIDIZATION OF HOMOLOGOUS SINGLE-STRANDED DNAs. IT INTERACTS WITH LEXA CAUSING ITS ACTIVATION AND LEADING TO ITS AUTOCATALYTIC CLEAVAGE.						
										CC	-!- SIMILARITY: BELONGS TO THE RECA FAMILY.						
										CC	-!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).						
										CC	-!- SIMILARITY: BELONGS TO THE RECA FAMILY.						
										CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).						
										CC	CC	CC	CC	CC	CC	CC	CC
										EMBL	L23135; AAA6702.1;						
										DR	HSSP; P03017; 2REB.						
										DR	InterPro; IPR00155;						
										DR	PFAM; PF00154; recf; 1.						
										DR	PRINTS; PRO00142; RECAB.						
										DR	PROSITE; PS00321; RECAB-1.						
										DR	DR						
										DR	PROSITE; PS50162; RECAB-2.						
										DR	DR						
										DR	PROSITE; PS50163; RECAB-3.						
										KW	DNA damage; DNA recombination; SOS response; ATP-binding; DNA-binding.						
										FT	NP_BIND_71						
										SQ	SEQUENCE 348 AA; 37912 MW; 311E9393089739ED CRC64;						
										DR	3.5%;						
										DR	Score 8;						
										DR	Best Local Similarity 100.0%;						
										DR	Pred. No. 2.1;						
										DR	Mismatches 0;						
										DR	Indels 0;						
										DR	Gaps 0;						
										DR	LEHPEMLE 115						
										DR	315 LEHPEMLE 322						
										Y4OA_RHISN	RESULT 4						
										ID	YAOA_RHISN						
										AC	STANDARD;						
										DT	P55586;						
										DT	01-NOV-1997 (Rel. 35, Created)						
										DT	01-NOV-1997 (Rel. 35, Last sequence update)						

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2001, 12:58:57 ; Search time 10.12 Seconds  
(without alignments)  
764.994 Million cell updates/sec

Title: US-09-441-723-1

Perfect score: 226

Sequence: 1 MGPLPRTVELFVLDLSPYSW.....AHLLGKWMGTPPPAVNRL 226

## Scoring table:

OLIGO Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues

Word size : 0

Total number of hits satisfying chosen parameters : 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	225	99.6	225	1	GTK1_HUMAN	O9y293 homo sapien
2	16	7.1	225	1	GTK1_RAT	F24473 ratus norv
3	8	3.5	248	1	RECA_AQUPY	P33542 aquifex pyr
4	8	3.5	593	1	Y4OA_RHISN	P55586 rhizobium s
5	8	3.5	814	1	SEFC_SALEN	P33388 salmonella
6	7	3.1	71	1	RL29_RICPR	O9zcr3 rickettsia
7	7	3.1	83	1	MULI_PSEAE	P11221 pseudomonas
8	7	3.1	131	1	PRO1_PHLPR	P35079 phleum prat
9	7	3.1	131	1	PRO2_PHLPR	P04650 phleum prat
10	7	3.1	131	1	PRO3_PHLPR	O24282 phleum prat
11	7	3.1	133	1	PROF_BETPV	P25816 betula verr
12	7	3.1	133	1	PROF_MERAN	O9g894 mercurialis
13	7	3.1	202	1	RS2_PYRHO	O59295 pyrococcus
14	7	3.1	204	1	TD53_HUMAN	O16890 homo sapien
15	7	3.1	204	1	TD53_MOUSE	O16890 homo sapien
16	7	3.1	354	1	YHLL_YEAST	P38780 saccharomyces
17	7	3.1	854	1	ODP2_PSEAE	O592371 mus musculu
18	7	3.1	547	1	DDR2_MOUSE	Q16832 homo sapien
19	7	3.1	855	1	DDR2_HUMAN	P38781 saccharomyces
20	7	3.1	883	1	YHLL_YEAST	P40558 homo sapien
21	7	3.1	902	1	GLR4_HUMAN	P19493 ratus norv
22	7	3.1	902	1	GLR4_RAT	O02455 saccharomyces
23	7	3.1	1875	1	MLPL_YEAST	Q9za04 rickettsia
24	6	2.7	67	1	Y158_RICPR	P00264 paracoccus
25	6	2.7	71	1	HPT1_PARSP	P81290 bacillus st
26	6	2.7	88	1	RS116_BACST	P56047 helicobacte
27	6	2.7	93	1	RL23_HELPY	P51386 porphyra pu
28	6	2.7	118	1	FTRC_PORPU	Q02366 bos taurus
29	6	2.7	127	1	NBAM_BOVIN	P56556 homo sapien
31	6	2.7	130	1	RS11_XANCP	O9z3e9 xanthomonas
32	6	2.7	131	1	MY07_HUMAN	O95214 homo sapien
33	6	2.7	132	1	RL14_HALHA	O24787 halobacteri

## ALIGNMENTS

RESULT 1

ID GTK1\_HUMAN STANDARD; PRT; 225 AA.

AC O9y203; Q9P1SA; DT 01-OCT-2000 (Rel. 40, Created)

DT 01-OCT-2000 (Rel. 40, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE GLUTATHIONE S-TRANSFERASE, MITOCHONDRIAL (EC 2.5.1.18) (GST 1.3-1.3)

DE GLUTATHIONE S-TRANSFERASE SUBUNIT 13) (GST CLASS-KAPPA) (HDCMDA7P).

GN GSTK1.

OS Homo sapiens (Human).

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI\_TAXID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=BLOOD;

RX MEDLINE=20493367; PubMed=11042152;

RA Zhang Q.-H., Ye M., Wu X.-Y., Ren S.-X., Zhao C.-J., Fu G.-W., Shen Y., Fan H.-Y., Lu G., Zhou M.-X., R., Han Z.-G., Zhang J.-W., Tao J., Huang Q.-H., Zhou J., Hu G.-X., Gu J., Chen S.-J., Chen Z.; "Cloning and functional analysis of cDNAs with open reading frames for 300 previously undefined genes expressed in CD34+ hematopoietic stem/progenitor cells." Genome Res. 10:1546-1560(2000).

RN [2]

RP SEQUENCE FROM N.A.

RA Zhao Z., Huang X., Li N., Zhu X., Cao X.; "A novel gene from human dendritic cell." Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.

CC -!- FUNCTION: MIGHT CONFER PROTECTION AGAINST GENOTOXIC AND CYTOTOXIC ELECTROPHILES IN THE MITOCHONDRIAL COMPARTMENT (BY SIMILARITY).

CC -!- CATALYTIC ACTIVITY: RX + R-S-GLUTATHIONE = HX + R-S-GLUTATHIONE.

CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. KAPPA FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and that this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

CC DR EMBL; AF070657; AND20963.1; -

DR EMBL; AF068387; AAF65506.1; -

KW Transferase; Mitochondrion.

FT INIT\_MET 0 BY SIMILARITY.

FT CONFLICT 178 G -> R (IN REF. 2).

FT CONFLICT 219 P -> S (IN REF. 2).

SQ SEQUENCE 225 AA; 25365 MW; PE9IA5EEFOB091 CRC64;

Query Match Score 225; DB 1; Pred. No. 1.1e-221; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1.1e-221;

Db	2	NDKLRLR	8		DR PROSITE; PS000913; BROKAR-LIPOPROTEIN; 1.
				KW Outer membrane; Lipoprotein; Signal.	
RESULT	7	MUL_PSEAE	STANDARD;	PRT; 83 AA.	
ID	MUL_PSEAE				
AC	P11221;				
DT	01-JUL-1989 (Rel. 11, Created)				
DT	01-JUL-1989 (Rel. 11, Last sequence update)				
DT	01-OCT-2000 (Rel. 40, Last annotation update)				
DE	MAJOR OUTER MEMBRANE LIPOPROTEIN PRECURSOR (MUREIN-LIPOPROTEIN)				
DE	(LIPOPROTEIN I).				
GN	OPIR OR PA2853.				
OS	Pseudomonas aeruginosa.				
OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;				
OX	NCBI_TaxID=287;				
RN	[1]	SEQUENCE FROM N.A.			
RP	SEQUENCE FROM N.A.				
RC	STRAIN=ISOLATE PA2;				
FX	MEDLINE=89313294; PubMed=2473376;				
RA	Cornelis P., Boula A., Belarbi A., Guyonvarch A., Kammerer B.,				
RA	Hannaert V., Hubert J.-C.;				
RT	"Cloning and analysis of the gene for the major outer membrane				
RT	lipoprotein from <i>Pseudomonas aeruginosa</i> ."				
RL	Mol. Microbiol. 3:421-428(1989).				
RN	[2]	SEQUENCE FROM N.A.			
RP	SEQUENCE FROM N.A.				
FX	MEDLINE=89327122; PubMed=2502533;				
RA	Duchene M., Barron C., Schweizer A., von Sprecht B.-U., Domdey H.;				
RA	" <i>Pseudomonas aeruginosa</i> outer membrane lipoprotein I gene: molecular				
RT	cloning, sequence, and expression in <i>Escherichia coli</i> ."				
RL	J. Bacteriol. 171:4130-4137(1989).				
RN	[3]	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PAO1;				
FX	MEDLINE=92268853; PubMed=1588307;				
RA	Saint-Onge A., Romeo F., Lebel P., Masson L., Brousseau R.,				
RA	"Specificity of the <i>Pseudomonas aeruginosa</i> PAO1 lipoprotein I gene as				
RT	a DNA probe and PCR target region within the <i>Pseudomonas aeruginosa</i> ."				
RL	J. Gen. Microbiol. 138:73-74(1992).				
RN	[4]	SEQUENCE FROM N.A.			
RP	SEQUENCE FROM N.A.				
RC	STRAIN=PAO1;				
FX	MEDLINE=20437337; PubMed=10984043;				
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,				
RA	Hickey M.J., Brinkman F.S.L., Huffnagle W.O., Kowalik D.J., Lagrou M.,				
RA	Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,				
RA	Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,				
RA	Spencer K.A., Spencer D.H., Wong G.K.S., Wu Z., Paulsen I.T.,				
RA	Reizer J., Sauer M.H., Hancock R.E.W., Lory S., Olson M.V.;				
RA	"Complete genome sequence of <i>Pseudomonas aeruginosa</i> PAO1, an				
RT	"opportunist pathogen."				
RL	Nature 406:959-964(2000).				
CC	-1- SURCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID ANCHOR.				
CC					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).				
CC					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).				
CC					
DR	EMBL; X13748; CAA32013_1; -				
DR	EMBL; M25761; AAA22880_1; -				
DR	EMBL; X58714; CAA01550_1; -				
DR	EMBL; A07695; CAA00707_1; -				
DR	EMBL; AE004712; AAC06241_1; -				
DR	PIR; S04834; S04834.				
DR	PIR; A33854; A33854.				
DR	PIR; A44834; A44834.				
DR	PROSITE; PS00392; PROFILIN.				
DR	PROSITE; PS00414; PROFILIN.				
DR	Actin-binding; Cytoskeleton; Multigene family.				
KW	Query Match 3.1%; Score 7; DB 1; Length 131;				
KW	Best Local Similarity 100.0%; Pred. No. 8.9;				
KW	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
SEQUENCE 131 AA; 14235 MW; 1B97A7CD339714C CRC64;					

RESULT		5			
ISEFC_SALEN		STANDARD;			
PAC	P33388;		PRY;	814	AA.
	01-FEB-1994	(Rel.	28,	Created)	
	01-FEB-1994	(Rel.	28,	Last sequence update)	
	01-FEB-1996	(Rel.	33,	Last annotation update)	
DDE	OUTER MEMBRANE USHER PROTEIN SEFC PRECURSOR.				
GN	SEFC.				
GN	Salmonella enteritidis.				
OS	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;				
OC	Salmonella.				
OX	NCB_TaxID=592;				
	{1}				
RN	SEQUENCE FROM N.A.				
RR	STRAIN=2765-38;				
RX	MEDLINE=91239677;				
RA	Clouthier S.C., Mueller K.-H., Doran J.L., Collinson S.K., Kay W.W.;				
RT	"Characterization of three fimbrial genes, sefABC, of Salmonella enteritidis.",				
RL	J. Bacteriol. 175:2533-2534(1993).				
CC	- FUNCTION: INVOLVED IN THE EXPORT AND ASSEMBLY OF THE SEFA FIMBRIAL SUBUNIT.				
CC	- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE (BY SIMILARITY).				
CC	-1- SIMILARITY: BELONGS TO THE FIMBRIAL EXPORT USHER FAMILY.				

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial/  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).

CC InterPro: IPR001865; -.  
 DR pfam: PF00318; Ribosomal\_S2; 1.

DR PROSITE: PS00395; RIBOSOMAL\_S2.

DR PROSITE: PS00962; RIBOSOMAL\_S2\_1; 1.

DR PROSITE: PS00963; RIBOSOMAL\_S2\_2; 1.

KW Ribosomal protein.  
 SQ SEQUENCE 202 AA; 23022 MW; 686F037D7AEFB57 CRC64;

Query Match 3.1%; Score 7; DB 1; Length 204;  
 Best Local Similarity 100.0%; Pred. No.13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 134 EPOSILA 140  
 DR 63 EPOSILA 69

RESULT 14  
 TPD53\_HUMAN STANDARD; PRT; 204 AA.  
 AC Q16590;  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE TUMOR PROTEIN D53 (HD53) (D52-LIKE 1).  
 GN TPDS2L1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NCBI\_TAXID=9606;  
 RN [1]

RP SEQUENCE FROM N.A.  
 TISSUE\_Breast carcinoma;  
 RX PubMed=97001154; Published=8812487;

RA Byrne J.A., Mattei M.-G., Basset P.;

RT "Definition of the tumor protein D52 (TPD52) gene family through cloning of D52 homologues in human (hd53) and mouse (md52)." ;

RL Genomics 35:523-532(1996).

CC -!- SUBUNIT: FORMS HOMODIMER OR HETERO DIMER WITH OTHER MEMBERS OF THE

CC FAMILY (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE TPD52 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - Eukaryotes. It forms part of the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----  
 DR AF004428; ANC98476.1; -.  
 DR MGII; MGII:1298386; Tpd52L1.

DR Coiled coil.  
 RT DOMAIN 22 73 COILED COIL (POTENTIAL),

SQ SEQUENCE 204 AA; 22515 MW; BAA49DAC7B7F6BE8 CRC64;

RP INTERACTIONS.  
 RX PubMed=9484778;

RA Byrne J.A., Nourse C.R., Basset P., Gunning P.;

RT "Identification of homo- and heteromeric interactions between members of the breast carcinoma-associated D52 protein family using the yeast two-hybrid system";

RL Oncogene 16:873-881(1998).

CC -!- SUBUNIT: FORMS HOMODIMER OR HETERO DIMER WITH OTHER MEMBERS OF THE

CC FAMILY.  
 CC -!- SIMILARITY: BELONGS TO THE TPD52 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR U44427; AAB0894.1; -.  
 DR EMBL: U44428; AAB0895.1; -.  
 DR MIM: 604069; -.  
 RW Coiled coil.  
 RT DOMAIN 22 73 COILED COIL (POTENTIAL),

SQ SEQUENCE 204 AA; 22449 MW; 6B3C336D5C0653C9 CRC64;

Query Match 3.1%; Score 7; DB 1; Length 204;  
 Best Local Similarity 100.0%; Pred. No.13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC InterPro: IPR001865; -.  
 DR pfam: PF00318; Ribosomal\_S2; 1.

DR PROSITE: PS00395; RIBOSOMAL\_S2.

DR PROSITE: PS00962; RIBOSOMAL\_S2\_1; 1.

DR PROSITE: PS00963; RIBOSOMAL\_S2\_2; 1.

KW Ribosomal protein.

SQ SEQUENCE 202 AA; 23022 MW; 686F037D7AEFB57 CRC64;

Query Match 3.1%; Score 7; DB 1; Length 204;  
 Best Local Similarity 100.0%; Pred. No.13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 151 QAQGLLE 157  
 DR 4 QAQGLLE 10

RESULT 15  
 TD53\_MOUSE STANDARD; PRT; 204 AA.

AC Q54818;  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)

DE TUMOR PROTEIN D53 (MD53) (D52-LIKE 1).

GN TPD52L1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI\_TAXID=10090;

OX NCBITaxonID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97001154; PubMed=8812487;

RA Byrne J.A., Mattei M.-G., Basset P.;

RT "Definition of the tumor protein D52 (TPD52) gene family through cloning of D52 homologues in human (hd53) and mouse (md52)." ;

RL Genomics 35:523-532(1996).

CC -!- SUBUNIT: FORMS HOMODIMER OR HETERO DIMER WITH OTHER MEMBERS OF THE

CC FAMILY (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE TPD52 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR U44427; AAB0894.1; -.  
 DR EMBL: U44428; AAB0895.1; -.  
 DR MIM: 604069; -.  
 RW Coiled coil.

RT DOMAIN 22 73 COILED COIL (POTENTIAL),

SQ SEQUENCE 204 AA; 22449 MW; 6B3C336D5C0653C9 CRC64;

Search completed: April 19, 2001, 13:00:52

Job time: 115 sec



Phleum pratense (Common timothy).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Phleum;  
 OC NCBI\_TaxID=15957;

[1]

RESULT 9  
 PRO2\_PHLPR STANDARD; PRT; 131 AA.

ID PRO2\_PHLPR  
 AC 024650;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE PROFILIN 2/4 (POLEN ALLERGEN PHL P 11) (PHL P XI).

GN PRO2 AND PRO4.  
 OS Phleum pratense (Common timothy).  
 OS Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Phleum.  
 OC NCBI\_TaxID=15957;

RN [1]

SEQUENCE FROM N.A.

RP TISSUE-Pollen; PubMed=9224949;  
 RX MEDLINE=9736182;

RA Asturias J.A., Arilla M.C., Bartolome B., Martinez A., Martinez A.,  
 Palacios R.;

RT "Sequence polymorphism and structural analysis of timothy grass  
 pollen profilin allergen (Phl p 11)." ;  
 RT pollen profilin allergen (Phl p 11);  
 RL Biochim. Biophys. Acta 1352:253-257(1997)

-!- FUNCTION: BINDS TO ACTIN AND AFFECTS THE STRUCTURE OF THE  
 CYTOSKELETON. AT HIGH CONCENTRATIONS, PROFILIN PREVENTS THE  
 POLYMERIZATION OF ACTIN, WHEREAS IT ENHANCES IT AT LOW  
 CONCENTRATIONS. BY BINDING TO PIP2, IT INHIBITS THE FORMATION OF  
 IP3 AND DG.

-!- SUBUNIT: OCCURS IN MANY KINDS OF CELLS AS A COMPLEX WITH MONOMERIC  
 ACTIN IN A 1:1 RATIO.

-!- DISEASE: CAUSE ALLERGIC REACTIONS.

-!- SIMILARITY: BELONGS TO THE PROFILIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 Sequence polymorphism and structural analysis of timothy grass  
 pollen profilin allergen (Phl p 11);  
 pollen profilin allergen (Phl p 11);  
 Biochim. Biophys. Acta 1352:253-257(1997)

-!- FUNCTION: BINDS TO ACTIN AND AFFECTS THE STRUCTURE OF THE  
 CYTOSKELETON. AT HIGH CONCENTRATIONS, PROFILIN PREVENTS THE  
 POLYMERIZATION OF ACTIN, WHEREAS IT ENHANCES IT AT LOW  
 CONCENTRATIONS. BY BINDING TO PIP2, IT INHIBITS THE FORMATION OF  
 IP3 AND DG.

-!- SUBUNIT: OCCURS IN MANY KINDS OF CELLS AS A COMPLEX WITH MONOMERIC  
 ACTIN IN A 1:1 RATIO.

-!- DISEASE: CAUSE ALLERGIC REACTIONS.

-!- SIMILARITY: BELONGS TO THE PROFILIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 or send an email to license@isb-sib.ch).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 or send an email to license@isb-sib.ch).

CC EMBL: Y09456; CAA70608\_1; -.  
 DR Y09458; CAA70610\_1; -.  
 DR InterPro: IPI002097; -.  
 DR Pfam: PF00235; profilin; 1.  
 DR PROSITE: PS00392; PROFILIN; 1.  
 DR PROSITE; PS00414; PROFILIN; 1.  
 DR Actin-binding; Cytoskeleton; Allergen; Multigene family.

CC SEQUENCE 131 AA; 14150 MW; 1B97B9DA084AE14C CRC64;

Query Match 3.1% Score 7; DB 1; Length 131;  
 Best Local Similarity 100.0%; Pred. No. 8.9;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 ITGIMKD 50  
 DR 47 ITGIMKD 53

RESULT 11  
 PROF\_BETVE STANDARD; PRT; 133 AA.

ID PROF\_BETVE  
 AC P25816;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE PROFILIN (POLEN ALLERGEN BET V 2) (BET V 11).  
 GN BETVII.  
 OS Betula verrucosa (White birch) (Betula pendula).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; euicicotyledons; core eudicots; Rosidae; eurosids I;  
 OC Fagales; Betulaceae; Betula.  
 OC NCBI\_TaxID=3505;

RN [1]

SEQUENCE FROM N.A.

RP TISSUE=Pollen;  
 RX MEDLINE=911313403; PubMed=1857985;

RA Valenta R., Duchene M., Pettenburger K., Sillaber C., Valent P.,  
 RA Bettelheim P., Breitenbach M., Rumpold H., Kraft D., Scheiner O.;

RT "Identification of profilin as a novel pollen allergen; IgE  
 autoimmunity in sensitized individuals.";  
 Science 253:557-560(1991).

RN [2]

RESULT 10  
 PRO2\_PHLPR STANDARD; PRT; 131 AA.

ID PRO2\_PHLPR  
 AC O2482;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE PROFILIN 3 (POLEN ALLERGEN PHL P 11) (PHL P XI).  
 DE PRO3.

RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).  
 RX MEDLINE=97169442; PubMed=9016715;  
 RA Fedorov A.A., Ball T., Mahoney N.M.; Valenta R., Almo S.C.;  
 RT "The molecular basis for allergen cross-reactivity: crystal structure  
 and IgE epitope mapping of birch pollen profilin."; crystal structure  
 structure 5:33:45 (1997).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).  
 RX MEDLINE=97276803; PubMed=9130496;  
 RA Fedorov A.A., Ball T., Valenta R., Almo S.C.;  
 RT "X-ray crystal structures of birch pollen profilin and Phl p 2.";  
 RL Int. Arch. Allergy Immunol. 113:109-113(1997).  
 RN [4]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=97415424; PubMed=9271223;  
 RA Schomberg D., Jockusch B.M.;  
 RT "Birch pollen profilin: structural organization and interaction with  
 poly-(L-proline) peptides as revealed by NMR.";  
 RL FEBS Lett. 411:291-295(1997).  
 CC -!- FUNCTION: BINDS TO ACTIN AND AFFECTS THE STRUCTURE OF THE  
 CC CYTOSKELETON. AT HIGH CONCENTRATIONS, PROFILIN PREVENTS THE  
 CC POLYMERIZATION OF ACTIN, WHEREAS IT ENHANCES IT AT LOW  
 CC CONCENTRATIONS. BY BINDING TO PIP2, IT INHIBITS THE FORMATION OF  
 CC IP3 AND DG. (BY SIMILARITY).  
 CC -!- SUBUNIT: OCCURS IN MANY KINDS OF CELLS AS A COMPLEX WITH MONOMERIC  
 CC ACTIN IN A 1:1 RATIO.  
 CC -!- DISEASE: CAUSE ALLERGIC REACTIONS.  
 CC -!- SIMILARITY: BELONGS TO THE PROFILIN FAMILY.  
 CC -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 EMBL; M65179; AAA16522.1; -.  
 CC -!- DISEASE: CAUSE ALLERGIC REACTIONS.  
 CC -!- SIMILARITY: BELONGS TO THE PROFILIN FAMILY.  
 CC -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR PIR; JC2082; JC2082;  
 DR PDB; 1CQA; 1-FBP-97.  
 DR InterPro; IPR00097; -.  
 DR Pfam; PF00235; Profilin; 1.  
 DR PRINTS; PRO0392; PROFILIN;  
 DR PROSITE; PS00414; PROFILIN; 1.  
 DR Actin-binding; Cytoskeleton; Allergen; 3D-structure.  
 KW SEQUENCE 133 AA; 14253 MW; 9443FC43786EE114A CRC64;  
 DR RS2\_PYRHO STANDARD; PRT; 202 AA.  
 DR RS2\_PYRHO STANDARD; PRT; 202 AA.  
 AC 059295;  
 AC -----  
 DT 30-MAY-2000 (Rel. 3.9, Created)  
 DT 01-OCT-2000 (Rel. 4.0, Last sequence update)  
 DT 01-OCT-2000 (Rel. 4.0, Last annotation update)  
 DE 30S RIBOSOMAL PROTEIN S2P.  
 GN RPS2 OR PH1629.  
 OS Pyrococcus horikoshii.  
 OC Archaea; Buryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus  
 OX NCBI\_TaxID=53953;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OT3;  
 RX MEDLINE=93344137; PubMed=9679194;  
 RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y., Nagai Y.,  
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,  
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfukui Y.,  
 RA Funahashi T., Tanaka T., Kudo Y., Yamazaki J., Kushida N., Oguchi A.,  
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
 RA Masuchi Y., Shizuya H., Kikuchi H.;  
 RT "Complete sequence and gene organization of the genome of a hyper-  
 thermophilic archaebacterium, Pyrococcus horikoshii OT3.";  
 RL DNA Res. 5:55-76 (1998).  
 CC -!- SIMILARITY: BELONGS TO THE S2P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR PIR; JC2082; JC2082;  
 DR PDB; 1CQA; 1-FBP-97.  
 DR InterPro; IPR00097; -.  
 DR Pfam; PF00235; Profilin; 1.  
 DR PRINTS; PRO0392; PROFILIN;  
 DR PROSITE; PS00414; PROFILIN; 1.  
 DR Actin-binding; Cytoskeleton; Allergen; 3D-structure.  
 KW SEQUENCE 133 AA; 14253 MW; 9443FC43786EE114A CRC64;  
 DR RS2\_PYRHO STANDARD; PRT; 202 AA.  
 DR RS2\_PYRHO STANDARD; PRT; 202 AA.  
 AC 049894;  
 AC -----  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DR 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE PROFILIN (ALLERGEN MER A 1).  
 OS Mercurialis annua (Annual mercury).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicots; core eudicots; Rosidae; eurosids I;  
 OC Malpighiales; Euphorbiaceae; Mercurialis.  
 OX NCBI\_TaxID=3986;  
 RN [1]  
 RP SEQUENCE FROM N.A.

Page 1

Gencore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

4 prc\_lein - protein search, using sw model

run on: April 19, 2001, 12:57:26 ; Search time 15.56 Seconds  
(w/o alignments)  
998.161 Million cell updates/sec

title:	US-09-441-733-1			
perfect score:	226			
sequence:	1 MGPLRPTVLFYDVLSPSW.....AHLGEKWMGPPIPPAVNARL 226			
scoring table:	ORIGO			
	Gapop 60.0 , Gapext 60.0			
searched:	198801 seqs, 68722935 residues			
word size :	0			
total number of hits satisfying chosen parameters: 198801				
minimum DB seq length:	0			
maximum DB seq length:	2000000000			
post-processing: Listing first 45 summaries				
database :	PIR_67;*			
	1: pirl;*			
	2: pir2;*			
	3: pir3;*			
	4: pir4;*			
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being Printed, and is derived by analysis of the total score distribution.				
SUMMARIES				
result No.	Score	Query Match Length	DB ID	Description
1	11	4.9	33	S17164 glutathione trans
2	8	3.5	77	C82988 hypothetical prote
3	8	3.5	305	T36056 hypothetical prote
4	8	3.5	305	T36056 recombinant prot
5	8	3.5	814	C40618 fibrillar outer mem
6	7	3.1	71	A71671 ribosomal protein
7	7	3.1	71	A71671 outer membrane lgp
8	7	3.1	83	F81873 probable membrane
9	7	3.1	95	profilin - common
10	7	3.1	131	JC2080 profilin - Europe
11	7	3.1	133	JC2082 probable chromosom
12	7	3.1	199	E75637 probable ribosomal
13	7	3.1	205	E71042 hypothesized protein
14	7	3.1	212	D70489 conserved hypoth
15	7	3.1	220	G83247 2-keto-3-deoxy-6-P
16	7	3.1	251	T08315 hypothetical prote
17	7	3.1	276	H81931 probable bis(5'-nu
18	7	3.1	279	T36082 probable binding P
19	7	3.1	287	T39197 yeast atp12 protei
20	7	3.1	288	D82271 conserved hypoth
21	7	3.1	347	T14313 hypothetical prote
22	7	3.1	357	T01434 NADPH HC toxin red
23	7	3.1	364	S7730 cbid protein - Syn
24	7	3.1	376	B75260 conserved hypoth
25	7	3.1	376	B8364 hypothetical prote
26	7	3.1	396	T27976 hypothetical prote
27	7	3.1	401	F83187 probable pyridoxal
28	7	3.1	458	H82679 hypothetical prote
29	7	3.1	498	H83078 two-component syst
30	7	3.1	547	H82679 dihydrolipoamidine a

ALIGNMENTS						
RESULT 1						
S17164	glutathione transferase (EC 2.5.1.18)	13 - rat				
C;Species:	Rattus norvegicus (Norway rat)					
C;Date:	21-Nov-1993	#sequence_revision	13-Mar-1997	#text_change	13-Mar-1997	
C;Accession:	S17164					
R;Harris, J.M.; Meyer, D.J.; Coles, B.; Ketterer, B.						
Biochem.	J. 278, 137-141, 1991					
A;Title:	A novel glutathione transferase (13-13)	isolated from the matrix of				
A;Reference number:	S17164	MUID:91354194				
A;Accession:	S17164					
A;Status:	preliminary					
A;Molecule type:	protein					
A;Residues:	1-33 <HAR>					
C;Keywords:	transferase					
RESULT 2						
C82988	hypothetical protein PA5271 [imported]	- Pseudomonas aeruginosa (strain PAO1				
C;Species:	Pseudomonas aeruginosa					
C;Date:	15-Sep-2010	#sequence_revision	15-Sep-2000	#text_change	31-Dec-2000	
C;Accession:	C82988					
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larsen, L.; Lory, S.; Olson, M.V.						
Nature	406, 959-964, 2000					
A;Title:	Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportu-					
A;Reference number:	A82950;	MUID:20437337				
A;Accession:	C82988					
A;Status:	preliminary					
A;Molecule type:	DNA					
A;Residues:	1-77 <STO>					
A;Cross-references:	GB:AE004939; GB:AE004091; MJD:99951575; PIDN:AAG08656.1;					
A;Experimental source:	strain PAO1					
C;Genetics:						
A;Gene:	PA5271					
Query Match Score 3.5%; DB 2; Length 77;						
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;						
Qy 9 ELFYDVLSPPS 19						
Db 8 ELFYDVLSPPS 18						
Query Match Score 4.9%; DB 2; Length 33;						
Best Local Similarity 100.0%; Pred. No. 0.00027; Mismatches 0; Indels 0; Gaps 0;						
Qy 9 ELFYDVLSPPS 19						
Db 8 ELFYDVLSPPS 18						

QY 149 AEQAQCLL 156  
Db 65 AEQAQCLL 72

RESULT 3  
T36056  
Species: Streptomyces coelicolor  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 18-Feb-2000.  
R;Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.  
Submitted to the EMBL Data Library, November 1998  
A;Reference number: Z21595  
A;Accession: T36056  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-305 <SAU>  
A;Cross-references: EMBL:AL034355; PIDN:CAA22222.1; GSPDB:GN00070; SCOEADB:SCD78.17c  
A;Experimental source: strain A3(2)  
C;Genetics:  
A;Gene: SCOEADB:SCD78.17c  
C;superfamily: Methanobacterium thermoautotrophicum conserved hypothetical protein MTH10

Query Match 3.5%; Score 8; DB 2; Length 305;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 LAAAEKAG 146  
Db 59 LAAAEKAG 66

RESULT 4  
A5020  
N;Alternate names: recombinase A  
C;Species: Aquifex pyrophilus  
C;Accession: A5020  
C;Database: 19-Mar-1997 sequence\_revision 19-Mar-1997 #text\_change 26-Aug-1999  
R;Wetmur, J.G.; Wong, D.M.; Ortiz, B.; Tong, J.; Reichert, F.; Gelfand, D.H.  
J. Biol. Chem. 269, 25928-25935, 1994  
A;Title: Cloning, sequencing, and expression of RecA proteins from three distantly related bacteria  
A;Reference number: A5020; MUID:5014407  
A;Accession: A5020  
A;Status: preliminary; nucleic acid sequence not shown  
A;Molecule type: DNA  
A;Residues: 1-348 <WET>  
A;Cross-references: GB:L23135; PIDN:9349396; PIDN:AAA67702.1; PID:9349397  
C;Superfamily: recombinase protein recA  
C;Keywords: ATP; DNA binding; DNA recombination; DNA repair; P-loop; SOS response  
F;71-78/Region: nucleotide-binding motif A (P-loop)  
F;145-150/Region: nucleotide-binding motif B  
F;77/Binding site: ATP (Lys) #status predicted

Query Match 3.5%; Score 8; DB 2; Length 318;  
Best Local Similarity 100.0%; Pred. No. 3.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 108 LEHPEMLE 115  
Db 315 LEHPEMLE 322

RESULT 5  
C40618  
Fimbrial outer membrane protein homolog SefC - *Salmonella enteritidis*  
C;Species: *Salmonella enteritidis*  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
R;Accession: C40618  
C;Reference number: S04834; MUID:89313294  
A;Accession: S04834  
A;Molecule type: DNA

J. Bacteriol. 175, 2523-2533, 1993  
A;Title: Characterization of three fimbrial genes, sefABC, of *Salmonella enteritidis*  
A;Reference number: A40618; MUID:33239677  
A;Contents: 27655-3b  
A;Accession: C40618  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-814 <CLO>  
A;Cross-references: GB:L11010; PIDN:9310649; PIDN:AAA27221.1; PID:9310650  
A;Note: sequence extracted from NCBI backbone (NCBIN:130387; NCBIP:130397)  
C;Superfamily: outer membrane usher protein fimD  
C;Keywords: membrane protein

Query Match 3.5%; Score 8; DB 1; Length 814;  
Best Local Similarity 100.0%; Pred. No. 7.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 166 NQLKETTE 173  
Db 60 NQLKETTE 67

RESULT 6  
A71671  
ribosomal protein L29 - *Rickettsia prowazekii*  
C;Species: *Rickettsia prowazekii*  
C;Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 03-Nov-2000  
C;Accession: A71671  
R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsm<sup>m</sup>  
Nature 396, 133-140, 1998  
A;Title: The genome sequence of *Rickettsia prowazekii* and the origin of mitochondrial DNA  
A;Reference number: A71670; MUID:99039499  
A;Accession: A71671  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-71 <AND>  
A;Cross-references: GB:AJ235272; GB:AJ235269; NID:93861033; PIDN:CAA15091.1; PID:ej  
A;Experimental source: strain Madrid E  
C;Genetics:  
A;Gene: rpmC; RP651

Query Match 3.1%; Score 7; DB 2; Length 71;  
Best Local Similarity 100.0%; Pred. No. 8.1;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 68 NDILKLRL 74  
Db 2 NDILKLRL 8

RESULT 7  
A33854  
outer membrane lipoprotein I precursor - *Pseudomonas aeruginosa*  
C;Species: *Pseudomonas aeruginosa*  
C;Date: 27-Feb-1990 #sequence\_revision 27-Feb-1990 #text\_change 31-Dec-2000  
C;Accession: A33854; S04834; A44834; A83288  
R;Duchene, M.; Barron, C.; Schweizer, A.; von Specht, B.U.; Dondey, H.  
J. Bacteriol. 171, 4130-4137, 1989  
A;Title: *Pseudomonas aeruginosa* outer membrane lipoprotein I gene: molecular cloning  
A;Reference number: A33854; MUID:89327122  
A;Accession: A33854  
A;Molecule type: DNA  
A;Residues: 1-83 <DUC>  
A;Cross-references: GB:M25761; NID:9151334; PIDN:AAA25880.1; PID:9151335  
R;Corneelis, P.; Bouia, A.; Belarbi, A.; Guyonvarch, A.; Kammerer, B.; Hannaert, V.;  
Mol. Microbiol. 3, 421-428, 1989  
A;Title: Cloning and analysis of the gene for the major outer membrane lipoprotein  
A;Reference number: S04834; MUID:89313294  
A;Accession: S04834  
A;Molecule type: DNA

A;Residues: 1-83 <CORS>  
A;Cross-references: EMBL:X13748; PIDN:CAA32013\_1; PID:945345  
A;Note: the authors translated the codon GAA for residue 78 as Gly  
R;Saint-Onge, A.; Romeyer, F.; Lebel, P.; Masson, L.; Brousseau, R.  
J. Gen. Microbiol. 138, 733-741, 1992  
A;Title: Specificity of the *Pseudomonas aeruginosa* PAO1 lipoprotein I gene as a DNA probe  
A;Reference number: A44834; MUID:322268853  
A;Accession: A44834  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-83 <SAI>  
A;Cross-references: GB:X58714; GB:S36066; NID:9433509; PIDN:CAA41550\_1; PID:9433510  
A;Experimental source: PAO1  
A;Note: sequence extracted from NCBI backbone (NCBIN:103666, NCBIPI:103667)  
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizouchi, S.D.; Warrener, P.; Hickey, M.J.; Bradian, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen  
A;Reference number: A82950; MUID:20437337  
A;Accession: A83288  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-83 <STO>  
A;Cross-references: GB:AE004712; GB:AE004091; NID:9948940; PIDN:AAG06241\_1; GSDB:GN001  
A;Experimental source: strain PAO1  
C;Genetics:  
A;Gene: opri; PA2853  
C;Keywords: lipid binding; lipoprotein; membrane protein  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-83/Product: lipoprotein I #status predicted <MAT>

Query Match Score 7; DB 2; Length 83;	Best Local Similarity 100.0%; Pred. No. 9.3;	Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;		

Qy 113 MLEKASR 119  
Db 76 MLEKASR 82

RESULT 8  
F81873  
Probable membrane protein NMA1081 [Imported] - *Neisseria meningitidis* (group A strain Z2  
C;Species: *Neisseria meningitidis*  
C;Accession: F81873  
R;Parkhill, J.; Achtman, M.; James, K.D.; Churcher, C.; Klee, S.R.; Morel, J.; Holroyd, S.; Jagels, K.; Leathier, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, Nature 404, 502-506, 2000  
A;Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.  
A;Reference number: A81775; MUID:20222556  
A;Accession: F81873  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-95 <PAR>  
A;Cross-references: GB:AL167755; GB:AL157959; NID:97379742; PIDN:CA884344\_1; PID:9737977  
A;Experimental source: serogroup A, strain Z2491  
C;Genetics:  
A;Gene: NMA1081

Query Match Score 7; DB 2; Length 95;	Best Local Similarity 100.0%; Pred. No. 11;	Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;		

Qy 44 ITGIMKD 50  
Db 59 ITGIMKD 65

RESULT 9  
E75637  
probable chromosome partitioning protein para family - *Deinococcus radiodurans* (strain 31-Mar-2000)  
C;Species: *Deinococcus radiodurans*  
C;Accession: E75637  
R;White, O.; Elsen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.; Shan, M.; Vaithayanathan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, S.; Smith, H.O.; Ventler, J.C.; Fraser, C.M.; Science 286, 1571-1577, 1999  
A;Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans*  
A;Reference number: A75250; MUID:20036896  
A;Accession: E75637  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-199 <WHI>

A;Cross-references: GB:AE001827; NID:96460959; PID:AAF12679.1; TIGR:DRC0025  
 A;Experimental source: strain RI  
 C;Genetics:  
 A;Gene: DRC0025  
 A;Map position: plasmid  
 A;Genome: plasmid  
 A;Note: Plasmid CP1

Query Match 3.1%; Score 7; DB 2; Length 199;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 7; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

Qy 37 LQLRPSL 43  
 Db 130 LQLRPSL 136

RESULT 12  
 E71042 probable ribosomal protein S2 - Pyrococcus horikoshii  
 C;Species: Pyrococcus horikoshii  
 C;Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 18-Aug-2000  
 C;Accession: E71042  
 R;Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Hino, Y.; Yamamoto, S.; Sekine, M.; Ohfukui, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushida, N.; Oguchi, A.;NRA Res. 5, 55-56, 1998  
 A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon  
 A;Reference number: A71000; MUID:98344137  
 A;Status: Preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-205 <RAW>  
 R;Cross-references: GB:AP000006; NID:93236133; PID:BA30741.1; PID:g3258058  
 A;Experimental source: strain OR3  
 A;Note: this accession replaces an interim accession for a sequence replaced by GenBank  
 C;Genetics:  
 A;Gene: PH1629  
 C;Superfamily: yeast ribosomal protein S1.e

Query Match 3.1%; Score 7; DB 2; Length 205;  
 Best Local Similarity 100.0%; Pred. No. 22;  
 Matches 7; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

Qy 134 EP0SILA 140  
 Db 66 EP0SILA 72

RESULT 13  
 D70489 hypothetical protein aq\_2203 - Aquifex aeolicus  
 C;Species: Aquifex aeolicus  
 C;Date: 08-MAY-1998 #sequence\_revision 08-May-1998 #text\_change 05-Nov-1999  
 C;Accession: D70489  
 R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov V.  
 Nature 392, 353-358, 1998  
 A;Title: The complete genome of the hyper-thermophilic bacterium Aquifex aeolicus.  
 A;Reference number: A70300; MUID:98196666  
 A;Accession: D70489  
 A;Status: Preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-212 <AQE>  
 A;Cross-references: GB:AE000777; NID:92984377; PID:AAC07907.1; PID:g2984389; GB:AE000655  
 A;Experimental source: strain VF5  
 C;Genetics:  
 A;Gene: aq\_2203  
 C;Superfamily: Aquifex aeolicus hypothetical protein aq\_2203

Query Match 3.1%; Score 7; DB 2; Length 212;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 7; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

Qy 204 ELLAHL 210  
 Db 64 ELLAHL 70

A;Cross-references: GB:AE001827; NID:96460959; PID:AAF12679.1; TIGR:DRC0025  
 A;Experimental source: strain RI  
 C;Genetics:  
 A;Gene: DRC0025  
 A;Map position: plasmid  
 A;Genome: plasmid  
 A;Note: Plasmid CP1

Query Match 3.1%; Score 7; DB 2; Length 199;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 7; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

Qy 92 LERGSLS 98  
 Db 156 LERGSLS 162

RESULT 14  
 G83247 2-keto-3-deoxy-6-phosphogluconate aldolase PA3181 [Imported] - Pseudomonas aeruginosa  
 C;Species: Pseudomonas aeruginosa  
 C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C;Accession: G83247  
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.; Adamian, S.; Yuan, Y.; Brody, L.W.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lory, S.; Olson, M.V.; Nature 406, 959-964, 2000  
 A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic  
 A;Reference number: A82250; MUID:20437337  
 A;Accession: G83247  
 A;Status: Preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-220 <STO>  
 A;Cross-references: GB:AE004742; GB:AE004091; NID:99949294; PIDN:AA06569.1; GSPDB  
 A;Experimental source: strain PAO1  
 C;Genetics:  
 A;Gene: PA3181  
 C;Superfamily: 2-dehydro-3-deoxyphosphogluconate aldolase

Query Match 3.1%; Score 7; DB 2; Length 220;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 7; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

Qy 140 AAAEKAG 146  
 Db 84 AAAEKAG 90

RESULT 15  
 T08315 hypothetical protein H1185 [Imported] - Halobacterium sp. (strain NRC-1) plasmid pC;Species: Halobacterium sp.  
 A;Variety: strain NRC-1  
 C;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 03-Nov-2000  
 C;Accession: T08315  
 R;Ng, W.V.; Clujo, S.A.; Smith, T.M.; Bumgarner, R.E.; Baskin, D.; Faust, J.; Hall, Genome Res. 8, 1131-1141, 1998  
 A;Title: Snapshot of a large dynamic replicon in a halophilic archaeon: megaplasmid A;Reference number: Z16408; MUID:99063795  
 A;Accession: T08315  
 A;Status: translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-251 <DAS>  
 A;Cross-references: EMBL:AF016485; NID:92822278; HALOSP:H1185  
 A;Experimental source: strain NRC-1  
 C;Genetics:  
 A;Gene: HALOSP:H1185  
 A;Genome: Plasmid pNRC100  
 C;Superfamily: Halobacterium plasmid pNRC100 hypothetical protein H1185

Query Match 3.1%; Score 7; DB 2; Length 251;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 7; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

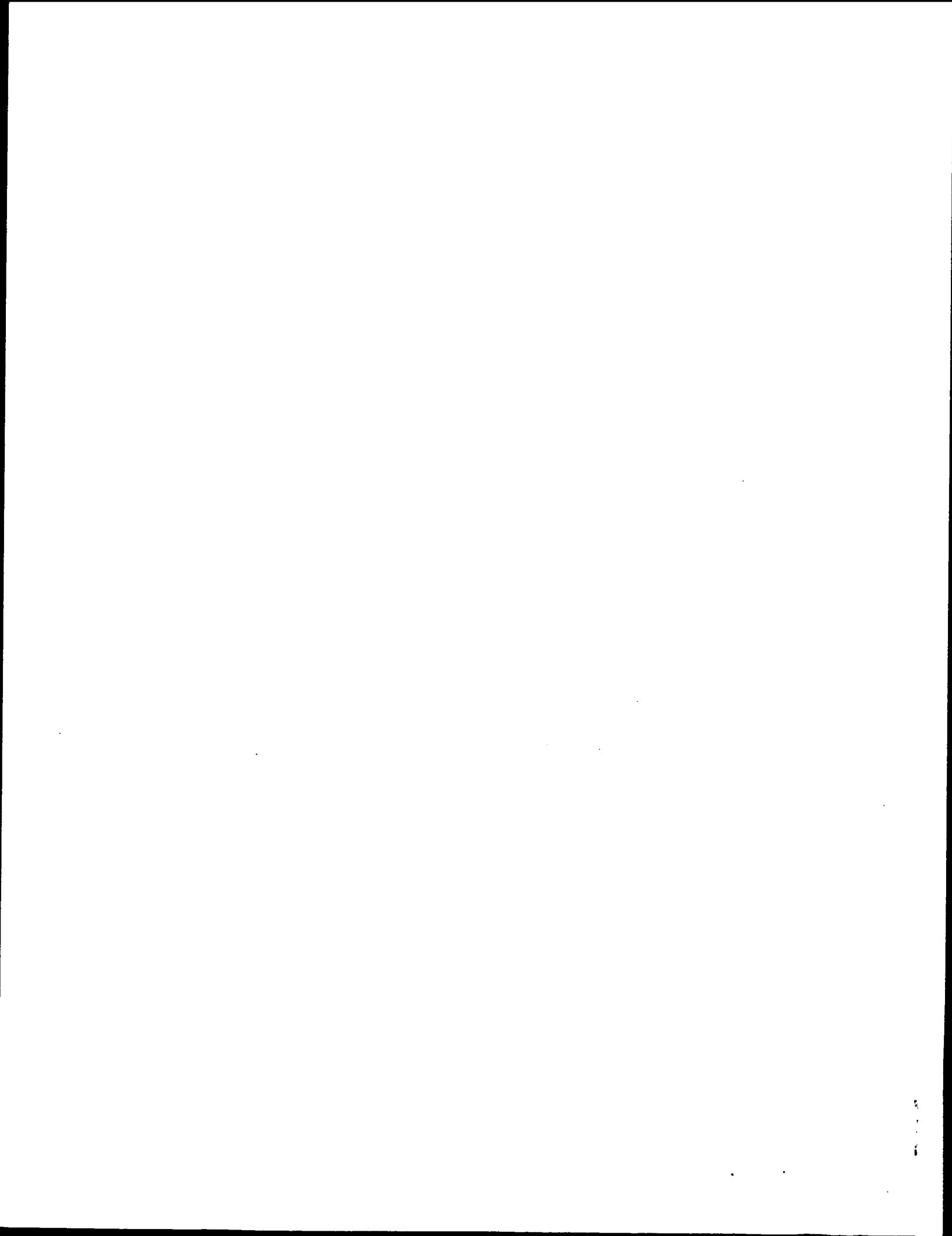
Qy 204 ELLAHL 210  
 Db 64 ELLAHL 70

Fri Apr 20 10:52:39 2001

us-09-441-723-1.rpr

Page 5

Search completed: April 19, 2001, 12:59:14  
Job time: 108 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2001, 12:57:21 ; Search time 14.05 Seconds  
(without alignments)  
309.014 Million cell updates/sec

Title: US-09-441-723-1  
Perfect score: 226  
Sequence: 1 MGPLPRTVELFYDVLSPYSW.....AHLLGEKWMGGPIPPAYNRL 226

Scoring table: OLIQO  
Gapop 60.0 , Gapext 60.0

Searched: 185757 seqs, 19210857 residues

Word size : 0

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_AA:  
 1: /cgn2\_6/ptodata/2/1aa/5A..COMB..pep:  
 2: /cgn2\_6/ptodata/2/1aa/5B..COMB..pep:  
 3: /cgn2\_6/ptodata/2/1aa/6A..COMB..pep:  
 4: /cgn2\_6/ptodata/2/1aa/6B..COMB..pep:  
 5: /cgn2\_6/ptodata/2/1aa/PCUTUS..COMB..pep:  
 6: /cgn2\_6/ptodata/2/1aa/backfiles1..pep:  
 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	226	100.0	226	3	US-09-978-174-1	Sequence 1, Appli
2	16	7.1	226	3	US-09-978-174-3	Sequence 3, Appli
3	8	3.5	814	1	US-08-533-788A-42	Sequence 42, Appli
4	7	3.1	63	2	US-08-577-447C-9	Sequence 9, Appli
5	7	3.1	131	1	US-07-846-932-11	Sequence 11, Appli
6	7	3.1	131	1	US-08-469-555-11	Sequence 11, Appli
7	7	3.1	133	1	US-07-846-992-3	Sequence 3, Appli
8	7	3.1	133	1	US-08-649-555-3	Sequence 1, Appli
9	7	3.1	204	2	US-08-715-204-1	Sequence 10, Appli
10	7	3.1	204	2	US-08-691-814B-10	Sequence 1, Appli
11	7	3.1	204	3	US-08-632-597-1	Sequence 1, Appli
12	7	3.1	214	2	US-08-672-447C-13	Sequence 13, Appli
13	7	3.1	221	3	US-09-531-563-22	Sequence 22, Appli
14	7	3.1	226	2	US-08-572-447C-15	Sequence 15, Appli
15	7	3.1	291	1	US-08-687-379-14	Sequence 14, Appli
16	7	3.1	291	1	US-08-687-379-16	Sequence 16, Appli
17	7	3.1	685	3	US-09-031-563-21	Sequence 21, Appli
18	7	3.1	854	2	US-08-056-647B-20	Sequence 20, Appli
19	7	3.1	854	2	US-08-237-401A-20	Sequence 4, Appli
20	7	3.1	855	1	US-08-336-347A-4	Sequence 8, Appli
21	7	3.1	902	1	US-07-718-575-8	Sequence 2, Appli
22	7	3.1	902	1	US-08-259-164-2	Sequence 8, Appli
23	7	3.1	902	1	US-08-681-206-8	Sequence 8, Appli
24	7	3.1	902	2	US-08-186-269A-8	Sequence 8, Appli
25	7	3.1	1055	3	US-09-031-563-27	Sequence 27, Appli
26	7	3.1	1315	3	US-09-031-563-2	Sequence 2, Appli
27	7	3.1	1315	3	US-09-031-563-25	Sequence 25, Appli

## ALIGNMENTS

RESULT 1  
US-09-978-174-1

; Sequence 1, Application US/08978174  
; Patent No. 603009  
; GENERAL INFORMATION:  
; APPLICANT: Shah, Purvi  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Tal, Preeti  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: NEW GLUTATHIONE-S-TRANSFERASE  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/978,174  
; FILING DATE: Herewith  
; CLASSIFICATION:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0430 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-845-4166  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 226 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: Single  
; TOPOLOGY: Linear  
; IMMEDIATE SOURCE:  
; LIBRARY: BLADTUT04  
; CLONE: 1554593  
; US-08-978-174-1

Query Match 100.0%; Score 226; DB 3; Length 226;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-223;  
 Matches 226; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

---

Qy 1 MGPLPRTVELFDVLSPIWGLPEILCYQNTINNQLRPLSLTGIMKDSGNKPGLLP 60  
 Db 1 MGPLPRTVELFDVLSPIWGLPEILCYQNTINNQLRPLSLTGIMKDSGNKPGLLP 60

Qy 61 RKGLYMANDLKLRLHHLQIPHPKDFLSVMLKGSSLAMARLTAVNLHEPMLEKASRE 120  
 Db 61 RKGLYMANDLKLRLHHLQIPHPKDFLSVMLKGSSLAMARLTAVNLHEPMLEKASRE 120

Qy 121 LMWRYWSRNEDTEPOSLAIAAEKAGNSAEGAQGLEKTATPKVKNQLEKTTAACRYGA 180  
 Db 121 LMWRYWSRNEDTEPOSLAIAAEKAGNSAEGAQGLEKTATPKVKNQLEKTTAACRYGA 180

Qy 181 FGLPTVAHVHDGQTHMLFGSDRMELLAHLGEKWMGPPIPPAVNRL 226  
 Db 181 FGLPTVAHVHDGQTHMLFGSDRMELLAHLGEKWMGPPIPPAVNRL 226

RESULT 3  
 US-08-233-788A-42  
 ; Sequence 42, Application US/08233788A  
 ; Patent No. 563617  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Doran, James L.  
 ; APPLICANT: Kay, William W.  
 ; APPLICANT: Clouthier, Sharon C.  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTION  
 ; NUMBER OF SEQUENCES: 61  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Seed and Berry  
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue  
 ; CITY: Seattle  
 ; STATE: Washington  
 ; COUNTRY: U.S.A.  
 ; ZIP: 98104-7092  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/233,788A  
 ; FILING DATE: 26-APR-1994  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: King, Joshua  
 ; REGISTRATION NUMBER: 35,570  
 ; REFERENCE/DOCKET NUMBER: 920043, 403C2  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (206) 622-4900  
 ; FAX: (206) 682-6031  
 ; TELEX: 3723816 SEEDANBERRY  
 ; INFORMATION FOR SEQ ID NO: 42:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 814 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-233-788A-42

---

US-08-978-174-3  
 ; Sequence 3, Application US/08978174  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shah, Purvi  
 ; APPLICANT: Hillman, Jennifer L.  
 ; APPLICANT: Lal, Preeti  
 ; APPLICANT: Corley, Neil C.  
 ; TITLE OF INVENTION: NEW GLUTATHIONE-S-TRANSFERASE  
 ; NUMBER OF SEQUENCES: 3  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 ; STREET: 3174 Porter Drive  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSEQ for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/978,174  
 ; FILING DATE: Herewith  
 ; CLASSIFICATION:  
 ; PRIORITY APPLICATION NUMBER:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Billings, Lucy J.  
 ; REGISTRATION NUMBER: 36,749  
 ; REFERENCE/DOCKET NUMBER: PE-0430 US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 650-555-0555  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 226 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: GenBank  
 ; CLONE: ?  
 ; US-08-978-174-3

Query Match 7.1%; Score 16; DB 3; Length 226;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-08;

---

Qy 166 NQLKETTE 173  
 Db 60 NQLKETTE 67

RESULT 4  
 US-08-572-447C-9  
 ; Sequence 9, Application US/08572447C  
 ; Patent No. 5955090  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Knapp, Bernhard  
 ; APPLICANT: Hungerer, Klaus-Dieter  
 ; APPLICANT: Broker, Michael  
 ; APPLICANT: Von Specht, Bernd-Ulrich  
 ; APPLICANT: Domdey, Horst  
 ; TITLE OF INVENTION: Immunogenic Hybrid Protein Oprf-Opri  
 ; Derived From Pseudomonas aeruginosa Membrane Proteins.

NUMBER OF SEQUENCES: 15  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Flanagan, Henderson, Farabow, Garrett &  
 ADDRESS: Dunner  
 STREET: 1300 I Street, N.W., Suite 700  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20005-3315

COMPUTER READABLE FORM:  
 MEDIUM TYPE: FLOPPY disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/846,992  
 FILING DATE: 1992-06-06  
 CLASSIFICATION: 435  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US 07/353,844  
 FILING DATE: 18-MAY-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Jones III, Harry C  
 REGISTRATION NUMBER: 20,280  
 REFERENCE/DOCKET NUMBER: 6530-011

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 863-9099  
 TELEFAX: (212) 863-8864/9741  
 TELEX: 66141 PENNTE 11:

INFORMATION FOR SEQ ID NO: 11:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 131 amino acids  
 TYPE: AMINO ACID  
 TOPOLOGY: Linear  
 MOLECULE TYPE: peptide  
 HYPOTHETICAL: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Phleum pratense

FEATURE:  
 OTHER INFORMATION: Amino acid identity with P14  
 OTHER INFORMATION: allergen from Betula verrucosa is 77%

US-08-572-447C-9

Query Match Similarity 3.1%; Score 7; DB 1; Length 131;  
 Best Local Similarity 100.0%; Pred. No. 11; Mismatches 0; Indels 0; Gaps 0;

Qy 113 MLEKASR 119  
 Db 56 MLEKASR 62

RESULT 6  
 US-08-469-555-11  
 Sequence 11, Application US/08469555

US-07-846-992-11  
 Sequence 11, Application US/07846992

PATENT NO. 5583046

GENERAL INFORMATION:  
 APPLICANT: Valentia, Rudolf  
 APPLICANT: Duchene, Michael  
 APPLICANT: Pettenthaler, Karin  
 APPLICANT: Breitenbach, Michael  
 APPLICANT: Kraft, Dietrich  
 APPLICANT: Rumpold, Helmut  
 APPLICANT: Scheiner, Otto  
 APPLICANT: Scheiner, Otto  
 TITLE OF INVENTION: Birch Pollen Allergen P14 for Diagnosis  
 TITLE OF INVENTION: and Therapy of Allergic Diseases  
 NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Pennie & Edmonds  
 STREET: 1155 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: U.S.A.  
 ZIP: 10036-2711

COMPUTER READABLE FORM:  
 MEDIUM TYPE: FLOPPY disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25

APPLICATION NUMBER: US/08/469,555  
 FILING DATE: 06-JUN-1995  
 CLASSIFICATION: 435  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US/07/846,992  
 FILING DATE: 06-JUN-1992

APPLICATION NUMBER: US/07/353, 844  
 FILING DATE: 18-MAY-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Jones III, Harry C  
 REGISTRATION NUMBER: 20, 280  
 REFERENCE/DOCKET NUMBER: 6530-011  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 790-9090  
 TELEX: (212) 869-8864/9741  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 133 amino acids  
 TYPE: AMINO ACID  
 TOPOLOGY: linear  
 MOLECULE TYPE: Peptide  
 HYPOTHETICAL: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Betula verrucosa  
 FEATURE: Amino acid sequence identity  
 OTHER INFORMATION: With profilin of other organisms is as follows:  
 OTHER INFORMATION: With human profilin, 28% with calf and mouse,  
 OTHER INFORMATION: 26% with yeast and 25% with Acanthamoeba  
 US-07-846-992-3

OTHER INFORMATION: Amino acid identity with P14  
 OTHER INFORMATION: Allergen from Betula verrucosa is 77%  
 US-08-469-555-11

Query Match 3.1%; Score 7; DB 1; Length 133;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 3.1%; Score 7; DB 1; Length 131;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 8  
 US-08-469-555-3

GENERAL INFORMATION:  
 Sequence 3, Application US/08469555  
 Patent No. 5648242

APPLICANT: Valenta, Rudolf  
 DUCHENE, Michael  
 PETTENBURGER, Karin  
 BREITENBACH, Michael  
 APPLICANT: Kraft, Dietrich  
 RUMPOLD, Helmut  
 SCHEINER, Otto  
 TITLE OF INVENTION: Birch Pollen Allergen P14 for Diagnosis  
 NUMBER OF SEQUENCES: 11  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Penne & Edmonds  
 STREET: 1155 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: U.S.A.  
 ZIP: 10036-2711

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/846, 992  
 FILING DATE: 1992-06-06  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/353, 844  
 FILING DATE: 18 MAY-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Jones III, Harry C  
 REGISTRATION NUMBER: 20, 280  
 REFERENCE/DOCKET NUMBER: 6530-011  
 TELECOMMUNICATION INFORMATION:

INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 133 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 HYPOTHETICAL: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Betula verrucosa  
 FEATURE: Amino acid sequence identity  
 OTHER INFORMATION: with profilin of other organisms is as follows:  
 OTHER INFORMATION: 30% with human profilin, 28% with calf and mouse,  
 OTHER INFORMATION: 26% with yeast and 25% with Acanthamoeba  
 US-08-469-555-3

RESULT 9  
 Query Match 3.1%; Score 7; DB 1; Length 133;  
 Best Local Similarity 100.0%; Pred. No. 11; Mismatches 0;  
 Matches 7; Conservative 0; Gaps 0;  
 Qy 44 ITGIMKD 50  
 Db 49 ITGIMKD 55

RESULT 10  
 Query Match 3.1%; Score 7; DB 2; Length 204;  
 Best Local Similarity 100.0%; Pred. No. 17; Mismatches 0;  
 Matches 7; Conservative 0; Gaps 0;  
 Qy 151 QAQGLE 157  
 Db 4 QAQGLE 10

US-08-715-204-1  
 Query Match 3.1%; Score 7; DB 2; Length 204;  
 Best Local Similarity 100.0%; Pred. No. 17; Mismatches 0;  
 Matches 7; Conservative 0; Gaps 0;  
 Qy 151 QAQGLE 157  
 Db 4 QAQGLE 10

RESULT 10  
 Sequence 10, Application US/08691814B  
 Patent No. 5981218  
 GENERAL INFORMATION:  
 APPLICANT: Rio, Marie-Christine  
 APPLICANT: Tomasetto, Catherine  
 APPLICANT: Basset, Paul  
 APPLICANT: Byrne, Jennifer  
 TITLE OF INVENTION: Isolated Nucleic Acid Molecules Useful  
 AS Leukemia Markers and in Breast Cancer Prognosis  
 NUMBER OF SEQUENCES: 124  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
 STREET: 1100 New York Ave., NW, Suite 600  
 CITY: Washington  
 STATE: DC  
 COUNTRY: USA  
 ZIP: 20005-3934  
 COMPUTER READABLE FORM:  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/691,814B  
 FILING DATE: 31-JUL-1996  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/002,183  
 FILING DATE: 09-AUG-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Steffe, Eric K.  
 REGISTRATION NUMBER: 36,688  
 REFERENCE/DOCKET NUMBER: 1383.0090001  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-371-2600  
 TELEFAX: 202-371-2543  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 204 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-691-814B-10

RESULT 11  
 Sequence 11, Application US/09162597  
 Patent No. 6043343  
 GENERAL INFORMATION:

APPLICANT: Bandman, Olga  
 APPLICANT: Au-Young, Janice  
 APPLICANT: Goli, Surya K  
 APPLICANT: Hillman, Jennifer.  
 APPLICANT: Zweiiger, Gary B.

TITLE OF INVENTION: A NOVEL TUMOR PROTEIN  
 NUMBER OF SEQUENCES: 7  
 CURRENT APPLICATION DATA:  
 ADDRESSSEES: Incyte Pharmaceuticals, Inc.  
 STREET: 3174 Porter Drive  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: U.S.  
 ZIP: 94304

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSEQ Version 1.5  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/162,597  
 FILING DATE:  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 08/715,204  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Billings, Lucy J.  
 REFERENCE/DOCKET NUMBER: PF-0126 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-555-0555  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 204 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 IMMEDIATE SOURCE:  
 LIBRARY:  
 CLONE: Consensus  
 US-09-162-597-1

Query Match 3.1%; Score 7; DB 3; Length 204;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 7; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

Qy 151 QAGGLE 157  
 Db 4 QAGGLE 10

RESULT 12  
 US-08-572-447C-13  
 .GENERAL INFORMATION:  
 ; Patent No. 5955090

APPLICANT: Knap, Bernhard  
 APPLICANT: Hungerer, Klaus-Dieter  
 APPLICANT: Broker, Michael  
 APPLICANT: Von Specht, Bernd-Ulrich  
 APPLICANT: Domdey, Horst

TITLE OF INVENTION: Immunogenic Hybrid Protein OpF-OprI  
 TITLE OF INVENTION: Derived From Pseudomonas aeruginosa Membrane Proteins.  
 NUMBER OF SEQUENCES: 15  
 CURRENT SEQUENCE ADDRESS:  
 ADDRESSSEE: Finnegan, Henderson, Farabow, Garrett & Dunner  
 STREET: 1300 I Street, N.W., Suite 700  
 CITY: Washington D.C.

RESULT 14  
 US-08-572-447C-15  
 ; Sequence 15, Application US/08572447C  
 ; Patent No. 5955090  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Knapp, Bernhard

COUNTRY: USA  
 ZIP: 20005-3315  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: FLOPPY disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/572,447C  
 FILING DATE: 14-DEC-1995  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: EP 94120023.0  
 FILING DATE: 16-DEC-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: M. Paul Barker  
 REGISTRATION NUMBER: 32,013  
 REFERENCE/DOCKET NUMBER: 05552.1395-00000  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 408-4000  
 TELEFAX: (202) 408-4400  
 INFORMATION FOR SEQ ID NO: 13:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 214 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-572-447C-13

Query Match 3.1%; Score 7; DB 2; Length 214;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 7; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

Qy 113 MLEKASR 119  
 Db 207 MLEKASR 213

RESULT 13  
 US-09-031-563-22  
 ; Sequence 22, Application US/09031563A  
 ; Patent No. 6022708  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Frederic de Sauvage  
 ; TITLE OF INVENTION: Fused  
 ; FILE REFERENCE: P1272  
 ; CURRENT APPLICATION NUMBER: US/09/031,563A  
 ; CURRENT FILING DATE: 1998-02-26  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SEQ ID NO: 22  
 ; LENGTH: 221  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-031-563-22

Query Match 3.1%; Score 7; DB 3; Length 221;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 7; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

Qy 38 QLRPSLI 44  
 Db 66 QLRPSLI 72

RESULT 14  
 US-08-572-447C-15  
 ; Sequence 15, Application US/08572447C  
 ; Patent No. 5955090  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Knapp, Bernhard

APPLICANT: Hungerer, Klaus-Dieter  
 APPLICANT: Broker, Michael  
 APPLICANT: Von Specht, Bernd-Ulrich  
 APPLICANT: Domdey, Horst  
 TITLE OF INVENTION: Immunogenic Hybrid Protein OprF-OporI  
 TITLE OF INVENTION: Derived From *Pseudomonas aeruginosa* Membrane Proteins.  
 NUMBER OF SEQUENCES: 15  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
 STREET: Dunner  
 CITY: 1300 I Street, N.W., Suite 700  
 STATE: Washington  
 COUNTRY: D.C.  
 ZIP: 20005-3315  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: FLOPPY disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US-08/572,447C  
 FILING DATE: 14-DEC-1995  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: EP 94120023.0  
 FILING DATE: 16-DEC-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: M. Paul Barker  
 REGISTRATION NUMBER: 32,013  
 REFERENCE/DOCKET NUMBER: 05552.1395-00000  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 408-4000  
 TELEFAX: (202) 408-4400  
 INFORMATION FOR SEQ ID NO: 15:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 226 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-572-447C-15

MEDIUM TYPE: Diskette-3.5 inch, 720 Kb storage  
 COMPUTER: IBM AT-compatible, 80286 Processor  
 OPERATING SYSTEM: MS-DOS version 6.0  
 SOFTWARE: WordPerfect version 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US-08/687,379  
 FILING DATE: 05-AUG-1996  
 CLASSIFICATION: 435  
 INFORMATION FOR SEQ ID NO: 14:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 291 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-687-379-14

Query Match	Match	Score	DB	Length
Qy	37	LQLRPSL	43	291
Db	132	LQLRPSL	138	291

Search completed: April 19, 2001, 12:58:52  
 Job time: 91 sec

---

RESULT 15  
 US-08-687-379-14  
 Sequence 14, Application US/08687379  
 GENERAL INFORMATION:  
 APPLICANT: Hoeger, Thomas  
 APPLICANT: Ultsch, Andreas  
 APPLICANT: Bach, Alfred  
 APPLICANT: Sterrer, Sylvia  
 APPLICANT: Lemaire, Hans-Georg  
 TITLE OF INVENTION: Subunits of Glutamate Receptors, Their Preparation and Their Use  
 NUMBER OF SEQUENCES: 26  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Keil & Weinkauf  
 STREET: 1101 Connecticut Avenue  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20036  
 COMPUTER READABLE FORM:



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: April 19, 2001, 12:58:37 ; Search time 24.77 Seconds

(without alignments)  
1069.398 Million cell updates/secTitle: US-09-441-723-1  
Perfect score: 226

1 MGPLPRTVELFYDVLSPSW.....AHLGEKWNNGPIPPAVNARL 226

Sequence: OLIIGO  
Gapop 60.0 , Gapext 60.0

Scoring table: 374700 seqs, 117207915 residues

Word size : 0

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

## Post-processing: Listing first 45 summaries

Database : SPTRMBL\_15:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_unclassified:\*
- 13: sp\_vertebrate:\*
- 14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	226	100.0	226	4	O9Y2Q3	homo sapien	09y2q3 hom sapien
2	178	78.8	226	4	Q9P1S4	homo sapien	Q9p1s4 hom sapien
3	8	3.5	283	4	Q9U1J2	homo sapien	Q9u1j2 hom sapien
4	8	3.5	305	2	Q9BY2	streptomyces	09by2 streptomyce
5	8	3.5	364	4	Q94959	homo sapien	Q94959 homo sapien
6	8	3.5	392	13	Q9TBFO	anguilla ja	Q9tbfo anguilla ja
7	8	3.5	419	13	Q9PUT7	brachydani	Q9put7 brachydani
8	8	3.5	426	13	Q9W611	cynops pyrr	Q9w611 cynops pyrr
9	7	3.1	83	2	Q85409	pseudomonas	085409 pseudomonas
10	7	3.1	83	2	Q85410	pseudomonas	085410 pseudomonas
11	7	3.1	83	2	Q85411	pseudomonas	085411 pseudomonas
12	7	3.1	83	2	Q85412	pseudomonas	085412 pseudomonas
13	7	3.1	83	2	Q85413	pseudomonas	085413 pseudomonas
14	7	3.1	83	2	Q85414	pseudomonas	085414 pseudomonas
15	7	3.1	83	2	Q85415	pseudomonas	085415 pseudomonas
16	7	3.1	83	2	Q85416	pseudomonas	085416 pseudomonas
17	7	3.1	83	2	Q85417	pseudomonas	085417 pseudomonas
18	7	3.1	83	2	Q85418	pseudomonas	085418 pseudomonas
19	7	3.1	83	2	Q85419	pseudomonas	085419 pseudomonas

20	7	3.1	83	2	Q85420	pseudomonas	085420 pseudomonas
21	7	3.1	83	2	Q85421	pseudomonas	085421 pseudomonas
22	7	3.1	83	2	Q85422	pseudomonas	085422 pseudomonas
23	7	3.1	83	2	Q85423	pseudomonas	085423 pseudomonas
24	7	3.1	83	2	Q85424	pseudomonas	085424 pseudomonas
25	7	3.1	83	2	Q85425	pseudomonas	085425 pseudomonas
26	7	3.1	83	2	Q85426	pseudomonas	085426 pseudomonas
27	7	3.1	83	2	Q85427	pseudomonas	085427 pseudomonas
28	7	3.1	83	2	Q85428	pseudomonas	085428 pseudomonas
29	7	3.1	83	2	Q85429	pseudomonas	085429 pseudomonas
30	7	3.1	83	2	Q85431	pseudomonas	085431 pseudomonas
31	7	3.1	83	2	Q85433	pseudomonas	085433 pseudomonas
32	7	3.1	83	2	Q85434	pseudomonas	085434 pseudomonas
33	7	3.1	83	2	Q85435	pseudomonas	085435 pseudomonas
34	7	3.1	83	2	Q85437	pseudomonas	085437 pseudomonas
35	7	3.1	83	2	Q85438	pseudomonas	085438 pseudomonas
36	7	3.1	83	2	Q85439	pseudomonas	085439 pseudomonas
37	7	3.1	83	2	Q85440	pseudomonas	085440 pseudomonas
38	7	3.1	83	2	Q85441	pseudomonas	085441 pseudomonas
39	7	3.1	83	2	Q85443	pseudomonas	085443 pseudomonas
40	7	3.1	83	2	Q85444	pseudomonas	085444 pseudomonas
41	7	3.1	83	2	Q85445	pseudomonas	085445 pseudomonas
42	7	3.1	95	2	Q9JUY6	neisseria m	Q9juy6 neisseria m
43	7	3.1	122	2	Q9RAN1	methyllobaci	Q9ran1 methyllobaci
44	7	3.1	189	2	Q9RHW6	borrelia bu	Q9rhw6 borrelia bu
45	7	3.1	192	2	Q9RHWS	borrelia bu	Q9rhs5 borrelia bu

## ALIGNMENTS

RESULT 1	09Y2Q3	PRELIMINARY;	PRT;	226 AA.
ID	Q9Y2Q3;			
AC	Q9Y2Q3;			
DT	01-NOV-1999 (TREMBLref. 12, Created)			
DT	01-NOV-1999 (TREMBLref. 12, Last sequence update)			
DT	01-NOV-1999 (TREMBLref. 12, Last annotation update)			
DE	GLUTATHIONE S-TRANSFERASE SUBUNIT 13 HOMOLOG.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
OX	NCBI_TaxID:9606;			
RN	[1]	SEQUENCE FROM N.A.		
RA	Mao M., Ye M., Zhang Q., Zhou J., Wu J., Shen Y., Kan L., He K., Gu B., Fu G., Chen S., Chen Z.;			
RA	"Human rGSTK1-1 homolog gene."			
RT	Submitted (JUN-1998) to the EMBL/GenBank/DDJB databases.			
RL	EMBL; AF076657; AAD20963.1;			
DR	KW Transferase.			
SQ	SEQUENCE 226 AA; 25497 MW; D3FDADF1D1533B5BAA CRC64;			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Query	Match	Length	DB	ID	Description
09y2q3	homo sapien	09y2q3 hom sapien			
Q9p1s4	homo sapien	Q9p1s4 hom sapien			
Q9u1j2	homo sapien	Q9u1j2 hom sapien			
Q9zby2	streptomyces	09zby2 streptomyce			
Q94959	homo sapien	Q94959 homo sapien			
Q9lbf0	anguilla ja	Q9lbf0 anguilla ja			
Q9put7	brachydani	Q9put7 brachydani			
Q9w611	cynops pyrr	Q9w611 cynops pyrr			
Q85409	pseudomonas	085409 pseudomonas			
Q85410	pseudomonas	085410 pseudomonas			
Q85411	pseudomonas	085411 pseudomonas			
Q85412	pseudomonas	085412 pseudomonas			
Q85413	pseudomonas	085413 pseudomonas			
Q85414	pseudomonas	085414 pseudomonas			
Q85415	pseudomonas	085415 pseudomonas			
Q85416	pseudomonas	085416 pseudomonas			
Q85417	pseudomonas	085417 pseudomonas			
Q85418	pseudomonas	085418 pseudomonas			
Q85419	pseudomonas	085419 pseudomonas			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Query	Match	Length	DB	ID	Description
09y2q3	homo sapien	09y2q3 hom sapien			
Q9p1s4	homo sapien	Q9p1s4 hom sapien			
Q9u1j2	homo sapien	Q9u1j2 hom sapien			
Q9zby2	streptomyces	09zby2 streptomyce			
Q94959	homo sapien	Q94959 homo sapien			
Q9lbf0	anguilla ja	Q9lbf0 anguilla ja			
Q9put7	brachydani	Q9put7 brachydani			
Q9w611	cynops pyrr	Q9w611 cynops pyrr			
Q85409	pseudomonas	085409 pseudomonas			
Q85410	pseudomonas	085410 pseudomonas			
Q85411	pseudomonas	085411 pseudomonas			
Q85412	pseudomonas	085412 pseudomonas			
Q85413	pseudomonas	085413 pseudomonas			
Q85414	pseudomonas	085414 pseudomonas			
Q85415	pseudomonas	085415 pseudomonas			
Q85416	pseudomonas	085416 pseudomonas			
Q85417	pseudomonas	085417 pseudomonas			
Q85418	pseudomonas	085418 pseudomonas			
Q85419	pseudomonas	085419 pseudomonas			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Query	Match	Length	DB	ID	Description
09y2q3	homo sapien	09y2q3 hom sapien			
Q9p1s4	homo sapien	Q9p1s4 hom sapien			
Q9u1j2	homo sapien	Q9u1j2 hom sapien			
Q9zby2	streptomyces	09zby2 streptomyce			
Q94959	homo sapien	Q94959 homo sapien			
Q9lbf0	anguilla ja	Q9lbf0 anguilla ja			
Q9put7	brachydani	Q9put7 brachydani			
Q9w611	cynops pyrr	Q9w611 cynops pyrr			
Q85409	pseudomonas	085409 pseudomonas			
Q85410	pseudomonas	085410 pseudomonas			
Q85411	pseudomonas	085411 pseudomonas			
Q85412	pseudomonas	085412 pseudomonas			
Q85413	pseudomonas	085413 pseudomonas			
Q85414	pseudomonas	085414 pseudomonas			
Q85415	pseudomonas	085415 pseudomonas			
Q85416	pseudomonas	085416 pseudomonas			
Q85417	pseudomonas	085417 pseudomonas			
Q85418	pseudomonas	085418 pseudomonas			
Q85419	pseudomonas	085419 pseudomonas			



DR	EMBL; AB020690; BAA74906.1; -.	PRINTS; PR00049; WILMSTUMOUR.
SO	SEQUENCE 364 AA; 41509 MW; 6E417AD96B3F0B93 CRC64;	DR PROSITE; PS00028; ZINC_FINGER_C2H2; 4.
Query Match	3.5%; Score 8; DB 4; Length 364;	DR SEQUENCE 419 AA; 46925 MW; 606ADFEDA619FECD CRC64;
Best Local Similarity	100.0%; Pred. No. 8.8;	SQ SEQUENCE
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Query Match 3.5%; Score 8; DB 13; Length 419;	
QY	204 ELLAHLG 211	Best Local Similarity 100.0%; Pred. No. 9.9;
Db	140 ELLAHLG 147	Mismatches 0; Indels 0; Gaps 0;
RESULT 6		
Q9IBFO	PRELIMINARY; PRT; 392 AA.	RESULT 8
ID Q9IBFO;		Q9W611 PRELIMINARY; PRT; 426 AA.
AC	01-OCT-2000 (TREMBLrel. 15, Created)	ID Q9W611
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)	AC Q9W611; PRELIMINARY;
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)	DT 01-NOV-1999 (TREMBLrel. 12, Created)
DE	WILMS / TUMOR PROTEIN.	DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
GN		DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
OS	Anguilla japonica (Japanese eel).	DE WT1.
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OS Cynops pyrrhogaster (Japanese common newt).
OC	Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillioidei;	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Anguillidae; Anguilla.	Amphibia; Batrachia; Caudata; Salamandroidea; Cynops.
OX	NCBI_TAXID=7937;	NCBI_TAXID=8330;
RN	[1]	RN [1]
RP	SEQUENCE FROM N.A.	RP SEQUENCE FROM N.A.
RC	TISSUE-KIDNEY;	RC TISSUE-TESTIS;
RA	Nakatsuru Y., Minami K., Yoshikawa A., Zhu J.J., Oda H., Masahito P.,	RA Nakayama Y., Yamamoto T., Matsuda Y., Abe S.I.;
RA	Ishikawa T.	RT "Cloning of cDNA for newt WT1 and the differential expression during spermatogenesis of the Japanese newt, Cynops pyrrhogaster";
RA	Okamoto N., Nakamura Y., Ishikawa T.	RT Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RT	"Bel WT1 sequence and expression in spontaneous nephroblastomas in Japanese eel."	RL EMBL: AB013888; BAA76399.1; -.
RT	Gene 245:245-251(2000); Gene 245:251-251(2000);	HSSP; P08046; 1AAY.
RL	EMBL; AB030741; BAA90558.1; -.	INTERPRO; IPR000822; -.
DR	SEQUENCE 392 AA; 43892 MW; DAE1A84828F43DF6 CRC64;	INTERPRO; IPR000976; -.
DR		PFAM; PF00096; zf-C2H2; 4.
DR		PRINTS; PR00049; WILMSTUMOUR.
DR		PROSITE; PS00028; ZINC_FINGER_C2H2; 4.
KW	Zinc-finger; Metal-binding; DNA-binding.	DR KW
SQ	SEQUENCE 426 AA; 47590 MW; 46A0B158A9F79C6C CRC64;	SQ SEQUENCE 426 AA; 47590 MW; 46A0B158A9F79C6C CRC64;
Query Match	3.5%; Score 8; DB 13; Length 392;	Query Match 3.5%; Score 8; DB 13; Length 426;
Best Local Similarity	100.0%; Pred. No. 9.3;	Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Mismatches 0; Indels 0; Gaps 0;	Mismatches 0; Indels 0; Gaps 0;
QY	175 ACRYGAFG 182	QY 175 ACRYGAFG 182
Db	111111 99	Db 90 ACRYGAFG 97
RESULT 7		RESULT 9
Q9PUT7	PRELIMINARY; PRT; 419 AA.	Q9PUT7 PRELIMINARY; PRT; 83 AA.
ID Q9PUT7;		ID 085409 PRELIMINARY;
AC	01-MAY-2000 (TREMBLrel. 13, Created)	AC 085409; PRELIMINARY;
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)	DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)	DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DE	WILMS / TUMOR SUPPRESSOR.	DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
OS	Brechydanio rerio (Zebrafish) (Zebra danio).	DE MAJOR OUTER MEMBRANE LIPOPROTEIN I (FRAGMENT).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OPRI.
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysii;	Pseudomonas mendocina.
OC	Cypriniformes; Cyprinidae; Rasborinae; Danio.	OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OX	NCBI_TAXID=7935;	OC Pseudomonas.
RN	SEQUENCE FROM N.A.	OX NCBI_TAXID=300;
RA	Smith S.I., Down M., Power M., Boyd A.W.; "Isolation and characterization of a cDNA encoding zebrafish (Danio rerio) WT-1."	RN [1]
RT	Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.	RP SEQUENCE FROM N.A.
RL		STRAIN=LMG 1223;
DR	HSSP; P08046; 1AAY.	RC MEDLINE#8931294; PubMed=2473376;
DR	INTERPRO; IPR000822; -.	RX Cornelis P., Boula A., Belarbi A., Guyonvarch A., Kammerer B.,
DR	INTERPRO; IPR000976; -.	RA Hannaert V., Hubert J.C.;
DR	PFAM; PF00096; zf-C2H2; 4.	RT "Cloning and analysis of the gene for the major outer membrane
DR	PFAM; PF02165; WT1; 3.	

RT lipoprotein from *Pseudomonas aeruginosa*.";  
 RL Mol. Microbiol. 3:421-428(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 STRAIN-LMG 1222;  
 RC De Vos D., Bouton C., Sarniguet A., De Vos P., Vauterin M.,  
 Cornelis P.;  
 "Sequence diversity of the oprI gene coding for the major outer  
 membrane lipoprotein I among rRNA group I pseudomonads.",  
 DR Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 KW Lipoprotein.  
 FT NON-TER 83 AA; 8843 MW; EF0F54D80801DBA1 CRC64;  
 SQ SEQUENCE 83 AA; 8843 MW; EF0F54D80801DBA1 CRC64;

Query Match 3.1%; Score 7; DB 2; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 113 MLEKASR 119  
 Db 76 MLEKASR 82

RESULT 10  
 ID 085410 PRELIMINARY; PRT; 83 AA.  
 AC 085410;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE MAJOR OUTER MEMBRANE LIPOPROTEIN I (FRAGMENT).  
 GN OPRI.  
 OS *Pseudomonas oleovorans*.  
 OC Bacteria; Proteobacteria; gamma subdivision; *Pseudomonadaceae*;  
 OC Pseudomonas.  
 NCBI-TAXID=301;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-LMG 2229;  
 RA Cornelis P., Bouia A., Belarbi A., Guyonvarch A., Kammerer B.,  
 Hannaert V., Hubert J.C.;  
 "Cloning and analysis of the gene for the major outer membrane  
 lipoprotein from *Pseudomonas aeruginosa*.",  
 RL Mol. Microbiol. 3:421-428(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-LMG 2229;  
 RA De Vos D., Bouton C., Sarniguet A., De Vos P., Vauterin M.,  
 Cornelis P.;  
 "Sequence diversity of the oprI gene coding for the major outer  
 membrane lipoprotein I among rRNA group I pseudomonads.",  
 DR Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 KW Lipoprotein.  
 FT NON-TER 83 AA; 8801 MW; EF0123A10676A2A1 CRC64;

AC 085411; 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE MAJOR OUTER MEMBRANE LIPOPROTEIN I (FRAGMENT).  
 GN OPRI.  
 OS *Pseudomonas pseudoalcaligenes*.  
 OC Bacteria; Proteobacteria; gamma subdivision; *Pseudomonadaceae*;  
 OC Pseudomonas.  
 NCBI-TAXID=330;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-LMG 6036;  
 RA Cornelis P., Bouia A., Belarbi A., Guyonvarch A., Kammerer B.,  
 Hannaert V., Hubert J.C.;  
 "Cloning and analysis of the gene for the major outer membrane  
 lipoprotein from *Pseudomonas aeruginosa*.",  
 RL Mol. Microbiol. 3:421-428(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-LMG 6036;  
 RA Cornelis P., Bouia A., Belarbi A., Guyonvarch A., Kammerer B.,  
 Hannaert V., Hubert J.C.;  
 "Sequence diversity of the oprI gene coding for the major outer  
 membrane lipoprotein I among rRNA group I pseudomonads.",  
 DR Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-LMG 6036;  
 RA Cornelis P., Bouia A., Belarbi A., Guyonvarch A., Kammerer B.,  
 Hannaert V., Hubert J.C.;  
 "Sequence diversity of the oprI gene coding for the major outer  
 membrane lipoprotein I among rRNA group I pseudomonads.",  
 DR Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-LMG 6036;  
 RA Cornelis P., Bouia A., Belarbi A., Guyonvarch A., Kammerer B.,  
 Hannaert V., Hubert J.C.;  
 "Cloning and analysis of the gene for the major outer membrane  
 lipoprotein from *Pseudomonas aeruginosa*.",  
 RL Mol. Microbiol. 3:421-428(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-LMG 1224;  
 RA De Vos D., Bouton C., Sarniguet A., De Vos P., Vauterin M.,  
 Cornelis P.;  
 "Sequence diversity of the oprI gene coding for the major outer  
 membrane lipoprotein I among rRNA group I pseudomonads.",  
 DR Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 RN [1]

RP SEQUENCE FROM N.A.  
 STRAIN=LMG 2257;  
 RX MEDLINE=89313294; PubMed=2473376;  
 RA Cornelis P., Bouia A., Belarbi A., Guyonvarch A., Kammerer B.,  
 RA Hannaert V., Hubert J.C.;  
 RT "Cloning and analysis of the gene for the major outer membrane  
 RT lipoprotein from *Pseudomonas aeruginosa*.";  
 RL Mol. Microbiol. 3:421-428(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 STRAIN=LMG 2257;  
 RC De Vos D., Bouton C., Sarniguet A., De Vos P., Vauterin M.,  
 RA Cornelis P.;  
 RT "Sequence diversity of the oprI gene coding for the major outer  
 membrane lipoprotein I among rRNA group I pseudomonads";  
 RT Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL: AF065953; AAD03508.1; -.  
 KW Lipoprotein.  
 FT NON\_TER 83 AA;  
 SQ 83 MW; FB5F4D8131BDEBE CRC64;  
 Query Match 3.1%; Score 7; DB 2; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Y 113 MLEKASR 119  
 |||||||  
 b 76 MLEKASR 82  
 b  
 RESULT 13  
 D 085413 PRELIMINARY; PRT; 83 AA.  
 C 085413 PRELIMINARY; PRT; 83 AA.  
 T 01-NOV-1998 (TREMBLrel. 08, Created)  
 T 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 T 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 T MAJOR OUTER MEMBRANE LIPOPOTINE I (FRAGMENT).  
 DE OPRI.  
 IN Pseudomonas pseudoalcaligenes.  
 JC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 NCBI\_TAXID=330;  
 XC  
 IN  
 AC  
 AP  
 SEQUENCE FROM N.A.  
 STRAIN=LMG 5516;  
 RX MEDLINE=89313294; PubMed=2473376;  
 RA Cornelis P., Bouia A., Belarbi A., Guyonvarch A., Kammerer B.,  
 RA Hannaert V., Hubert J.C.;  
 RT "Cloning and analysis of the gene for the major outer membrane  
 RT lipoprotein from *Pseudomonas aeruginosa*.";  
 RL Mol. Microbiol. 3:421-428(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC  
 RA De Vos D., Bouton C., Sarniguet A., De Vos P., Vauterin M.,  
 RA Cornelis P.;  
 RT "Sequence diversity of the oprI gene coding for the major outer  
 RT membrane lipoprotein I among rRNA group I pseudomonads";  
 RT Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL: AF065952; AAD03507.1; -.  
 KW Lipoprotein.  
 FT  
 SQ 83 AA; 8829 MW; E80854DF0F01DBA1 CRC64;  
 Query Match 3.1%; Score 7; DB 2; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Y 113 MLEKASR 119  
 |||||||  
 b 76 MLEKASR 82  
 b  
 RESULT 14  
 ID 085414 PRELIMINARY; PRT; 83 AA.  
 AC  
 DR 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE MAJOR OUTER MEMBRANE LIPOPOTINE I (FRAGMENT).  
 GN OPRI.  
 OS Pseudomonas putida.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 RN NCBI\_TAXID=303;  
 XC  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 STRAIN=LMG 2162;  
 RC De Vos D., Bouton C., Sarniguet A., De Vos P., Vauterin M.,  
 RA Cornelis P.;  
 RT "Sequence diversity of the oprI gene coding for the major outer  
 RT membrane lipoprotein I among rRNA group I pseudomonads";  
 RT Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.  
 RL EMBL: AF065954; AAD03509.1; -.  
 KW Lipoprotein.  
 FT  
 SQ 83 AA; 8827 MW; FB5F436EC8860DBBA CRC64;  
 Query Match 3.1%; Score 7; DB 2; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Y 113 MLEKASR 119  
 |||||||

Fri Apr 20 10:52:41 2001

us-09-441-723-1.rspt

Page 6

Db 76 MLEKASR 82

Search completed: April 19, 2001, 13:00:36  
Job time: 119 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: April 19, 2001, 12:58:12 ; search time 4.15 Seconds

(without alignments)  
137.860 Million cell updates/sec

Title: US-09-441-723-1

Perfect score: 226

Sequence: 1 MGPLPRTVBLFYDVLSPYSW.....AHLLGKWMGPIPPAVNARL 226

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 12259 seqs, 2531507 residues

Word size :

Total number of hits satisfying chosen parameters: 12259

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

## Post-processing: Listing first 45 summaries

Database : Pending\_Patents\_AA\_New \*

1: /cgn2\_6/ptodata/1/paa/pcr/\_new\_comb.pep:\*

2: /cgn2\_6/ptodata/1/paa/us07\_new\_comb.pep:\*

3: /cgn2\_6/ptodata/1/paa/us07\_new\_comb.pep:\*

4: /cgn2\_6/ptodata/1/paa/us08\_new\_comb.pep:\*

5: /cgn2\_6/ptodata/1/paa/us60\_new\_comb.pep:\*

6: /cgn2\_6/ptodata/1/paa/us60\_new\_comb.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	%	Score	DB 5;	Length 261;
Best Local Similarity									
					Matches 7; Conservative	3.18;	Score 7;	DB 5;	Length 261;
1	7	3.1	261	5	US-09-739-449-11148	Sequence 11148, A	Qy 140 AAAEKAG 146	Prod. No. 2.6;	Length 261;
2	7	3.1	688	6	US-60-248-505-865	Sequence 865, APP	Db 50 AAAEKAG 56	0; Indels 0;	Gaps 0;
3	6	2.7	62	5	US-09-739-449-10531	Sequence 10531, A			
4	6	2.7	161	5	US-09-739-449-12044	Sequence 12044, A			
5	6	2.7	207	5	US-09-739-449-9529	Sequence 9529, AP			
6	6	2.7	215	5	US-09-739-449-9343	Sequence 9343, AP			
7	6	2.7	235	5	US-09-739-449-9873	Sequence 9873, AP			
8	6	2.7	235	5	US-09-739-449-10784	Sequence 10784, A			
9	6	2.7	245	5	US-09-739-449-10257	Sequence 10257, A			
10	6	2.7	258	5	US-09-739-449-10586	Sequence 10586, A			
11	6	2.7	305	5	US-09-739-449-9486	Sequence 9486, AP			
12	6	2.7	315	5	US-09-739-449-11093	Sequence 11093, A			
13	6	2.7	325	5	US-09-739-449-12418	Sequence 12418, A			
14	6	2.7	327	5	US-09-739-449-13257	Sequence 13257, A			
15	6	2.7	329	5	US-09-739-449-12591	Sequence 12591, A			
16	6	2.7	337	5	US-09-739-449-11560	Sequence 11560, A			
17	6	2.7	338	5	US-09-739-449-12951	Sequence 12951, A			
18	6	2.7	447	5	US-09-739-449-11196	Sequence 11196, A			
19	6	2.7	456	5	US-09-739-449-13245	Sequence 13245, A			
20	6	2.7	474	5	US-09-739-449-11298	Sequence 11298, A			
21	6	2.7	496	6	US-60-248-505-1273	Sequence 1273, APP			
22	6	2.7	570	6	US-60-248-505-980	Sequence 980, APP			
23	6	2.7	730	6	US-60-248-505-921	Sequence 921, APP			
24	6	2.7	926	5	US-09-442-187	Sequence 187, APP			
25	6	2.7	926	5	US-09-421-124-187	Sequence 187, APP			
26	6	2.7	1116	5	US-09-739-449-11359	Sequence 11359, A			
27	6	2.7	2882	6	US-60-248-505-1095	Sequence 1095, AP			

## ALIGNMENTS

RESULT 1	US-09-739-449-11148	; Sequence 11148, Application US/09739449
		; GENERAL INFORMATION:
		; APPLICANT: Hinkle, Gregory J.
		; INVENTOR: Slater, Steven C.
		; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
		; FILE REFERENCE: 38-10154901C
		; CURRENT APPLICATION NUMBER: US/09-739-449-11782
		; CURRENT FILING DATE: 2000-12-19
		; PRIOR APPLICATION NUMBER: US 09/514,000
		; PRIOR FILING DATE: 2000-02-23
		; NUMBER OF SEQ ID NOS: 13351
		; SEQ ID NO: 11148
		; LENGTH: 261
		; TYPE: PRT
		; ORGANISM: Agrobacterium tumefaciens
		US-09-739-449-11148

Query Match	Match	Local Similarity	Best	Mismatches	Score	DB 5;	Length
1	7	3.18;	100.0%	0;	100.0%	261;	261;

Query Match	Match	Local Similarity	Best	Mismatches	Score	DB 5;	Length
Qy 140 AAAEKAG 146	Db 50 AAAEKAG 56	100.0%	0;	0;	100.0%	261;	261;

RESULT 2	US-09-248-505-865	; Sequence 865, Application US/60248505
		; GENERAL INFORMATION:
		; APPLICANT: Beasley, Ellen
		; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
		; INVENTION: PROTEINS, AND USES THEREOF
		; FILE REFERENCE: C1000918
		; CURRENT APPLICATION NUMBER: US/60/248,505
		; CURRENT FILING DATE: 2000-11-15
		; NUMBER OF SEQ ID NOS: 1998
		; SOFTWARE: FastSEQ for Windows Version 4.0
		; SEQ ID NO: 865
		; LENGTH: 688
		; TYPE: PRT
		; ORGANISM: Human
		US-60-248-505-865

Query Match 3.1%; Score 7; DB 6; Length 688

Best Local Similarity 100.0%; Pred. No. 6;  
Matches 7; Conservative 0; Mismatches 0;  
Indels 0; Gaps 0;

RESULT 3  
US-09-739-449-10531  
; Sequence 10531, Application US/09739449  
; GENERAL INFORMATION:  
; APPLICANT: Hinkle, Gregory J.  
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15490)C  
; CURRENT FILING DATE: 2000-12-19  
; PRIORITY APPLICATION NUMBER: US 09/739,449  
; PRIOR FILING DATE: 2000-02-23  
; NUMBER OF SEQ ID NOS: 13351  
; SEQ ID NO 10531  
; LENGTH: 62  
; TYPE: PRT  
; ORGANISM: Agrobacterium tumefaciens  
US-09-739-449-10531

Query Match 2.7%; Score 6; DB 5;  
Best Local Similarity 100.0%; Pred. No. 7.6;  
Matches 6; Conservative 0; Mismatches 0;  
Indels 0; Gaps 0;

Qy 167 QLKETT 172  
Db 5 QLKETT 10

RESULT 4  
US-09-739-449-12044  
; Sequence 12044, Application US/09739449  
; GENERAL INFORMATION:  
; APPLICANT: Hinkle, Gregory J.  
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15490)C  
; CURRENT FILING DATE: 2000-12-19  
; PRIORITY APPLICATION NUMBER: US 09/739,449  
; PRIOR FILING DATE: 2000-02-23  
; NUMBER OF SEQ ID NOS: 13351  
; SEQ ID NO 12044  
; LENGTH: 161  
; TYPE: PRT  
; ORGANISM: Agrobacterium tumefaciens  
US-09-739-449-12044

Query Match 2.7%; Score 6; DB 5;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 6; Conservative 0; Mismatches 0;  
Indels 0; Gaps 0;

Qy 104 TAVNLE 109  
Db 153 TAVNLE 158

RESULT 5  
US-09-739-449-9529  
; Sequence 9529, Application US/09739449  
; GENERAL INFORMATION:  
; APPLICANT: Hinkle, Gregory J.  
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15490)C  
; CURRENT APPLICATION NUMBER: US/09/739,449  
; PRIORITY FILING DATE: 2000-12-19  
; NUMBER OF SEQ ID NOS: 13351  
; SEQ ID NO 9529  
; LENGTH: 207  
; TYPE: PRT  
; ORGANISM: Agrobacterium tumefaciens  
US-09-739-449-9529

Query Match 2.7%; Score 6; DB 5;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 6; Conservative 0; Mismatches 0;  
Indels 0; Gaps 0;

Qy 138 ILAAAE 143  
Db 15 ILAAAE 20

RESULT 6  
US-09-739-449-9343  
; Sequence 9343, Application US/09739449  
; GENERAL INFORMATION:  
; APPLICANT: Hinkle, Gregory J.  
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15490)C  
; CURRENT APPLICATION NUMBER: US/09/739,449  
; CURRENT FILING DATE: 2000-12-19  
; PRIORITY APPLICATION NUMBER: US 09/514,000  
; PRIOR FILING DATE: 2000-02-23  
; NUMBER OF SEQ ID NOS: 13351  
; SEQ ID NO 9343  
; LENGTH: 215  
; TYPE: PRT  
; ORGANISM: Agrobacterium tumefaciens  
US-09-739-449-9343

Query Match 2.7%; Score 6; DB 5;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 6; Conservative 0; Mismatches 0;  
Indels 0; Gaps 0;

Qy 139 LAAEAK 144  
Db 159 LAAEAK 164

RESULT 7  
US-09-739-449-9873  
; Sequence 9873, Application US/09739449  
; GENERAL INFORMATION:  
; APPLICANT: Hinkle, Gregory J.  
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15490)C  
; CURRENT APPLICATION NUMBER: US/09/739,449  
; CURRENT FILING DATE: 2000-12-19  
; PRIORITY APPLICATION NUMBER: US 09/514,000  
; PRIOR FILING DATE: 2000-02-23  
; NUMBER OF SEQ ID NOS: 13351  
; SEQ ID NO 9873  
; LENGTH: 235  
; TYPE: PRT  
; ORGANISM: Agrobacterium tumefaciens  
US-09-739-449-9873

Query Match 2.7%; Score 6; DB 5;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 6; Conservative 0; Mismatches 0;  
Indels 0; Gaps 0;

```

; RESULT 8
; US-09-739-449-10784
; Sequence 10784, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO: 10784
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-10784

; RESULT 8
; Sequence 10784, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO: 10784
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-10784

Query Match 2.7%; Score 6; DB 5; Length 235;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO: 10784
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-10784

Query Match 2.7%; Score 6; DB 5; Length 235;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO: 10786
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-10786

Query Match 2.7%; Score 6; DB 5; Length 258;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO: 10586
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-10586

; RESULT 11
; US-09-739-449-9486
; Sequence 9486, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO: 9486
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-9486

Query Match 2.7%; Score 6; DB 5; Length 305;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO: 9486
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-9486

Query Match 2.7%; Score 6; DB 5; Length 305;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO: 9486
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-9486

; RESULT 12
; US-09-739-449-11093
; Sequence 11093, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO: 11093
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-11093

Query Match 2.7%; Score 6; DB 5; Length 245;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO: 11093
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-11093

; RESULT 10
; US-09-739-449-10586
; Sequence 10586, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO: 10586
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-10586

Query Match 2.7%; Score 6; DB 5; Length 315;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO: 10586
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-10586

```

```

CURRENT FILING DATE: 2000-12-19
; PRIORITY APPLICATION NUMBER: US 09/514,000
; PRIORITY FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO: 12591
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-12591

RESULT 13
US-09-739-449-12418
; Sequence 12418 ; Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; ATTORNEY: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO: 12418
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-12418

Query Match 2.7%; Score 6; DB 5; Length 329;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 141 AAEKAG 146
Db 98 AAEKAG 103

Search completed: April 19, 2001, 13:00:08
Job time: 116 sec

RESULT 14
US-09-739-449-13257
; Sequence 13257 ; Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; ATTORNEY: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO: 13257
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-13257

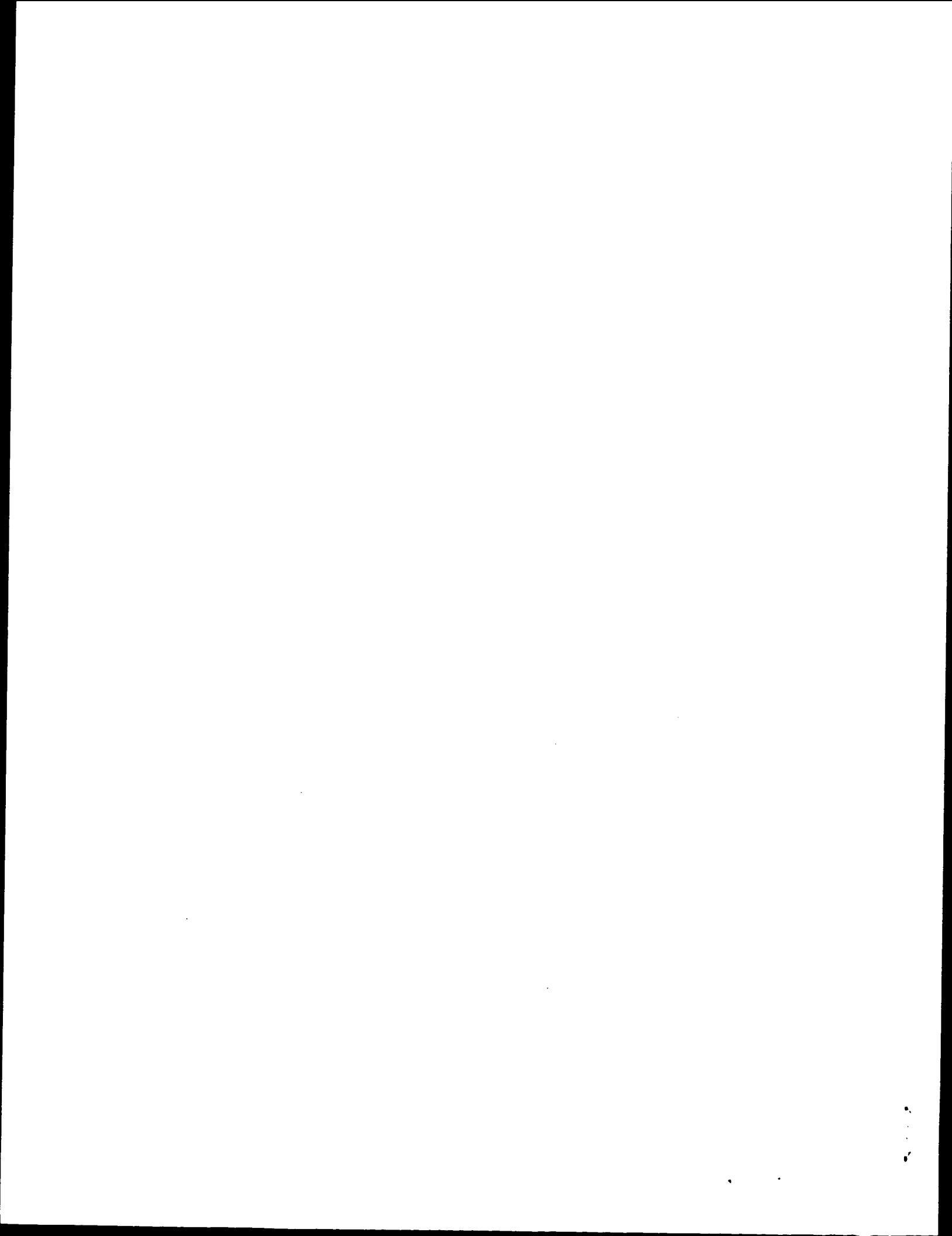
Query Match 2.7%; Score 6; DB 5; Length 327;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 153 QGLLEK 158
Db 59 QGLLEK 64

Search completed: April 19, 2001, 13:00:08
Job time: 116 sec

RESULT 15
US-09-739-449-12591
; Sequence 12591 ; Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; ATTORNEY: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449

```





GenCore version 4.5  
 Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model  
 Run on: April 19, 2001, 12:58:02 ; Search time 44.47 Seconds  
 (without alignments)  
 817.475 Million cell updates/sec

Title: US-09-441-723-1  
 Perfect score:  
 Sequence: 1 MGDLPRIVTELFYDVLSPYSW.....AHLLGEKWMGPIPPAVNARL 226

Scoring table: OLIGO  
 Gapop 60.0 , Gapext 60.0

Searched: 1009251 seqs, 160854530 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1009251

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :	Pending_Patents_AA_Main:*
1:	/con2_6/podata/2/paa/pcrus_comb.pep:*
2:	/cgn2_6/podata/2/paa/us06_comb.pep:*
3:	/cgn2_6/podata/2/paa/us07_comb.pep:*
4:	/cgn2_6/podata/2/paa/us080_comb.pep:*
5:	/cgn2_6/podata/2/paa/us081_comb.pep:*
6:	/cgn2_6/podata/2/paa/us082_comb.pep:*
7:	/cgn2_6/podata/2/paa/us083_comb.pep:*
8:	/cgn2_6/podata/2/paa/us084_comb.pep:*
9:	/cgn2_6/podata/2/paa/us085_comb.pep:*
10:	/cgn2_6/podata/2/paa/us086_comb.pep:*
11:	/con2_6/podata/2/paa/us087_comb.pep:*
12:	/cgn2_6/podata/2/paa/us088_comb.pep:*
13:	/cgn2_6/podata/2/paa/us089_comb.pep:*
14:	/cgn2_6/podata/2/paa/us090_comb.pep:*
15:	/cgn2_6/podata/2/paa/us091_comb.pep:*
16:	/cgn2_6/podata/2/paa/us092_comb.pep:*
17:	/cgn2_6/podata/2/paa/us093_comb.pep:*
18:	/cgn2_6/podata/2/paa/us094_comb.pep:*
19:	/cgn2_6/podata/2/paa/us095_comb.pep:*
20:	/cgn2_6/podata/2/paa/us096_comb.pep:*
21:	/cgn2_6/podata/2/paa/us097_comb.pep:*
22:	/cgn2_6/podata/2/paa/us098_comb.pep:*
23:	/cgn2_6/podata/2/paa/us60_comb.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## ALIGNMENTS

RESULT 1  
 US-09-441-723-1  
 Sequence 1, Application US/09441723  
 GENERAL INFORMATION  
 APPLICANT: Shah, Purvi L.  
 APPLICANT: Hillman, Jennifer L.  
 APPLICANT: Lal, Preeti  
 APPLICANT: Corley, Neil C.  
 TITLE OF INVENTION: NEW GLUTATHIONE-S-TRANSFERASE  
 NUMBER OF SEQUENCES: 3  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 STREET: 3174 Porter Drive  
 CITY: Palo Alto  
 STATE: CA  
 COUNTY: USA  
 ZIP: 94304

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEQ FOR Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/441,723  
 FILING DATE:  
 CLASSIFICATION:  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 08/978,174  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Billings, Lucy J.  
 REGISTRATION NUMBER: 36,749  
 REFERENCE/DOCKET NUMBER: PF-0430 US

## SUMMARIES

Result No.	Score	Query Match	Leng	DB ID	Description
1	226	100.0	226	18	US-09-441-723-1
2	150	66.4	150	23	US-60-147-499-4666
3	97	42.9	97	1	PCT-US00-06824176
4	92	40.7	120	23	US-60-245-201-153
5	92	40.7	120	23	US-60-245-201-226
6	16	7.1	226	18	US-69-441-723-3
7	12	5.3	91	1	PCT-US00-06824175
8	8	3.5	173	16	US-09-252-991A-23967
9	8	3.5	179	1	PCT-US00-05918-529
10	8	3.5			

TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 650-855-0555  
 ; TELEFAX: 650-845-4166  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 226 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: BLADUT04  
 ; CLONE: 1554593  
 ; US-09-441-723-1

RESULT 3  
 PCT-US00-06824-176  
 ; Sequence 176, Application PC/TUS0006824  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Human Genome Sciences, Inc.  
 ; TITLE OF INVENTION: 47 Human Secreted Proteins  
 ; FILE REFERENCE: PS511PCT  
 ; CURRENT APPLICATION NUMBER: PCT/US00/06824  
 ; CURRENT FILING DATE: 2000-03-16  
 ; EARLIER APPLICATION NUMBER: 60/125,359  
 ; EARLIER FILING DATE: 1999-03-19  
 ; EARLIER APPLICATION NUMBER: 60/168,664  
 ; EARLIER FILING DATE: 1999-12-03  
 ; NUMBER OF SEQ ID NOS: 180  
 ; SOFTWARE: Patentin Ver. 2.0  
 ; SEQ ID NO 176  
 ; LENGTH: 97  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 PCT-US00-06824-176

Query Match 100.0%; Score 226; DB 18; Length 226;  
 Best Local Similarity 100.0%; Pred. No. 9.8e-224;  
 Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGPLPRTVLFYDVLSPYSWLGFEILCQYQNTWNINLQRPSLTGIMKDSGNKPFGLLP 60  
 Db 1 MGPLPRTVLFYDVLSPYSWLGFEILCQYQNTWNINLQRPSLTGIMKDSGNKPFGLLP 60

Qy 61 RKGLYMANDKLRRHLQIPHFPKDFLSVMLRGSLSAMRFILTAVNLHPEMLESARE 120  
 Db 61 RKGLYMANDKLRRHLQIPHFPKDFLSVMLRGSLSAMRFILTAVNLHPEMLESARE 120

Qy 121 LMRVWSRNEDITEQSI1AAEAKGMSAEQAGLLEKIAITPKVKNQLKETTEAACRYGA 180  
 Db 121 LMRVWSRNEDITEQSI1AAEAKGMSAEQAGLLEKIAITPKVKNQLKETTEAACRYGA 180

Qy 181 FGLPITVAHVGDQTHMLFGSDRMELLAHLIGKEWMGPPIPNAVRL 226  
 Db 181 FGLPITVAHVGDQTHMLFGSDRMELLAHLIGKEWMGPPIPNAVRL 226

RESULT 2  
 US-60-147-499-4666  
 ; Sequence 4666, Application US/60147499  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dumas Milne Edwards, J.B.  
 ; ATTORNEY: Robert, S.  
 ; APPLICANT: Giordano, J.Y.  
 ; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
 ; FILE REFERENCE: GENSET.05APR2  
 ; CURRENT APPLICATION NUMBER: US/60/147,499  
 ; CURRENT FILING DATE: 1999-08-05  
 ; NUMBER OF SEQ ID NOS: 1935  
 ; SOFTWARE: Patent .pm  
 ; SEQ ID NO 4666  
 ; LENGTH: 150  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-60-147-499-4666

Query Match 66.4%; Score 150; DB 23; Length 150;  
 Best Local Similarity 100.0%; Pred. No. 8.8e-146;  
 Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGPLPRTVLFYDVLSPYSWLGFEILCQYQNTWNINLQRPSLTGIMKDSGNKPFGLLP 60  
 Db 1 MGPLPRTVLFYDVLSPYSWLGFEILCQYQNTWNINLQRPSLTGIMKDSGNKPFGLLP 60

Qy 61 RKGLYMANDKLRRHLQIPHFPKDFLSVMLRGSLSAMRFILTAVNLHPEMLESARE 120  
 Db 61 RKGLYMANDKLRRHLQIPHFPKDFLSVMLRGSLSAMRFILTAVNLHPEMLESARE 120

Qy 121 LMRVWSRNEDITEQSI1AAEAKGMSAEQAGLLEKIAITPKVKNQLKETTEAACRYGA 150  
 Db 121 LMRVWSRNEDITEQSI1AAEAKGMSAEQAGLLEKIAITPKVKNQLKETTEAACRYGA 150

RESULT 4  
 US-60-245-201-153  
 ; Sequence 153, Application US/60245201  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beasley, Ellen  
 ; TITLE OF INVENTION: ISOLATED HUMAN PHASE II DRUG-METABOLIZING PROTEINS, NUCLEIC ACID MOLECULES ENCODING THE SAME, AND USES THEREOF  
 ; TITLE OF INVENTION: HUMAN DRUG-METABOLIZING PHASE II PROTEINS, AND USES THEREOF  
 ; FILE REFERENCE: CLJ000879  
 ; CURRENT APPLICATION NUMBER: US/60/245,201  
 ; CURRENT FILING DATE: 2000-11-03  
 ; NUMBER OF SEQ ID NOS: 381  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 153  
 ; LENGTH: 120  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 US-60-245-201-153

Query Match 40.7%; Score 92; DB 23; Length 120;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-86;  
 Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 135 PQS11AAEAKGMSAEQAGLLEKIAITPKVKNQLKETTEAACRYGA 194  
 Db 29 PQS11AAEAKGMSAEQAGLLEKIAITPKVKNQLKETTEAACRYGA 194

Qy 195 HMFQSDRMELLAHLIGKEWMGPPIPNAVRL 226  
 Db 89 HMFQSDRMELLAHLIGKEWMGPPIPNAVRL 226

RESULT 5  
 US-60-245-201-226  
 ; Sequence 226, Application US/60245201  
 ; GENERAL INFORMATION:

APPLICANT: Beasley, Ellen  
 TITLE OF INVENTION: ISOLATED HUMAN PHASE II, NUCLEIC ACID MOLECULES ENCODING DRUG-METABOLIZING PROTEINS, NUCLEIC ACID MOLECULES ENCODING DRUG-METABOLIZING PHASE II PROTEINS, AND USES THEREOF  
 FILE REFERENCE: CL000879  
 CURRENT APPLICATION NUMBER: US/60/245,201  
 NUMBER OF SEQ ID NOS: 381  
 SEQ ID NO: 226  
 LENGTH: 120  
 TYPE: PRT  
 ORGANISM: Human  
 US-60-245-201-226

Query Match 40.7%; Score 92; DB 23; Length 120;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-86;  
 Matches 92; Conservative 0; Indels 0; Gaps 0;

Qy 135 POSTLAAAEATAGMSAEQAQGLLEKATPKVKNQLKETTEAACRYGAAGLPLTVAHVDQQT 194  
 Db 29 POSTLAAAEATAGMSAEQAQGLLEKATPKVKNQLKETTEAACRYGAAGLPLTVAHVDQQT 88

Qy 195 HMLFGSDRMELLAHLLGEKNGMGPPIPPAVNRL 226  
 Db 89 HMLFGSDRMELLAHLLGEKNGMGPPIPPAVNRL 120

RESULT 6  
 Sequence 3, Application US/09441723  
 GENERAL INFORMATION:  
 APPLICANT: Shah, Purvi  
 APPLICANT: Hilman, Jennifer L.  
 APPLICANT: Lal, Preeti  
 APPLICANT: Corley, Neill C.  
 TITLE OF INVENTION: NEW GLUTATHIONE-S-TRANSFERASE  
 NUMBER OF SEQUENCES: 3  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 STREET: 3174 Porter Drive  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastS2O for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/441,723  
 FILING DATE:  
 CLASSIFICATION:  
 APPLICATION NUMBER: 08/978-174  
 PRIORITY APPLICATION NUMBER:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Billings, Lucy J.  
 REGISTRATION NUMBER: 36,749  
 REFERENCE/DOCKET NUMBER: PF-0430 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 650-855-0555  
 TELEFAX: 650-845-4166

INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 226 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:

---

LIBRARY: GenBank  
 CLONE: ?  
 US-09-441-723-3

Query Match 7.1%; Score 16; DB 18; Length 226;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-08;  
 Matches 16; Conservative 0; Indels 0; Gaps 0;

Qy 9 ELFYDVLSPYSMLGFE 24  
 Db 9 ELFYDVLSPYSMLGFE 24

RESULT 7  
 PCT-US00-06824-175  
 Sequence 175, Application PC/TUS0006824  
 GENERAL INFORMATION:  
 APPLICANT: Human Genome Sciences, Inc.  
 TITLE OF INVENTION: Human Secreted Proteins  
 FILE REFERENCE: PSS1.PCT  
 CURRENT APPLICATION NUMBER: PCT/TUS00/06824  
 CURRENT FILING DATE: 2000-03-16  
 EARLIER APPLICATION NUMBER: 60/125,359  
 EARLIER FILING DATE: 1999-03-19  
 EARLIER APPLICATION NUMBER: 60/168,664  
 EARLIER FILING DATE: 1999-12-03  
 NUMBER OF SEQ ID NOS: 180  
 SOFTWARE: Patentin Ver. 2.0  
 SEQ ID NO 175  
 LENGTH: 97  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 PCT-US00-06824-175

Query Match 5.3%; Score 12; DB 1; Length 97;  
 Best Local Similarity 100.0%; Pred. No. 0.0004;  
 Matches 12; Conservative 0; Indels 0; Gaps 0;

Qy 196 MLFGSDRMELLA 207  
 Db 67 MLFGSDRMELLA 78

RESULT 8  
 US-09-252-991A-23967  
 Sequence 23967, Application US/09252991A  
 GENERAL INFORMATION:  
 APPLICANT: Rubenfield et al.  
 TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMYCETOGENOUS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: 107196.136  
 CURRENT FILING DATE: 1999-02-18  
 PRIOR APPLICATION NUMBER: US 60/074,788  
 PRIOR FILING DATE: 1998-02-18  
 NUMBER OF SEQ ID NOS: 33142  
 SEQ ID NO 23967  
 LENGTH: 91  
 TYPE: PRT  
 ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-23967

Query Match 3.5%; Score 8; DB 16; Length 91;  
 Best Local Similarity 100.0%; Pred. No. 4.9;  
 Matches 8; Conservative 0; Indels 0; Gaps 0;

Qy 149 AEQAQGLL 156  
 |||||||

```

Db 79 AEQAQGLL 86
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 3.0
RESULT 9
; Sequence 7934, Application US/09489039A
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709_2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; SEQ ID NO: 7934
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7934

Query Match 3 5%; Score 8; DB 18; Length 173;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 139 LAAAEKAG 146
Db 120 LAAAEKAG 127

RESULT 10
FILE REFERENCE: PCT-US00-05918-529
; Sequence 529, Application PC/TUS0005918
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Human Lung Cancer Associated Gene Sequences and Polypeptides
; FILE REFERENCE: PA104PCT
; CURRENT APPLICATION NUMBER: PCT/US00/05918
; CURRENT FILING DATE: 2000-03-08
; EARLIER APPLICATION NUMBER: 60/124,270
; EARLIER FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 529
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US00-05918-529

Query Match 3 5%; Score 8; DB 1; Length 179;
Best Local Similarity 100.0%; Pred. No. 9.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 96 SLSAMRFL 103
Db 25 SLSAMRFL 32

RESULT 11
; Sequence 7, Application US/09189527
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma Antibodies
; FILE REFERENCE: SL98-01
; CURRENT APPLICATION NUMBER: US/09/189,527
; CURRENT FILING DATE: 1998-11-10
; SEQ ID NO: 7
; LENGTH: 195
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-189-527-7

Query Match 3 5%; Score 8; DB 15; Length 195;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 204 ELLAHILG 211
Db 139 ELLAHILG 146

RESULT 12
; Sequence 7, Application US/09189527A
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma Antibodies
; FILE REFERENCE: SL98-01
; CURRENT APPLICATION NUMBER: US/09/189,527A
; CURRENT FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO: 7
; LENGTH: 195
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-189-527-7

Query Match 3 5%; Score 8; DB 15; Length 195;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 204 ELLAHILG 211
Db 139 ELLAHILG 146

RESULT 13
; Sequence 40, Application US/08233642A
; GENERAL INFORMATION:
; APPLICANT: Kay, William W.
; APPLICANT: Collinson, S. Karen
; APPLICANT: Clouthier, Sharon C.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SALMONELLA-
; NUMBER OF SEQUENCES: 58
; NUMBER OF SEQUENCES: -
; CORRESPONDENCE ADDRESS:
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS DOS
; SOFTWARE: PatentIn Release #1.0, version #1.25
; CURRENT APPLICATION DATA:
US-08-233-642A-40

```

APPLICATION NUMBER: US/08/233,642A  
 FILING DATE: 26-APR-1994  
 CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
 NAME: King, Joshua  
 REGISTRATION NUMBER: 35,570  
 REFERENCE/DOCKET NUMBER: 920043.403C3  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-4900  
 TELEFAX: (206) 682-6031  
 TELEX: 3723B36 SEEDANBERRY  
 INFORMATION FOR SEQ ID NO: 40:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 814 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein

US-08-233-642A-40

Query Match Score 8; DB 6; Length 814;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 166 NOLKETTE 173  
 Db 60 NOLKETTE 67

RESULT 14  
 US-60-186-662-1013  
 GENERAL INFORMATION:  
 APPLICANT: Bonazzi, Vivien  
 TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,  
 NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS,  
 TITLE OF INVENTION: AND USES THEREOF  
 FILE REFERENCE: CL000314  
 CURRENT APPLICATION NUMBER: US/60/186,662  
 CURRENT FILING DATE: 2000-03-03  
 NUMBER OF SEQ ID NOS: 1080  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO 1013  
 LENGTH: 46  
 TYPE: PRT  
 ORGANISM: HUMAN

US-60-186-662-1013

Query Match Score 8; DB 8; Length 55;  
 Best Local Similarity 100.0%; Pred. No. 33;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 95 GSLSAMR 101  
 Db 42 GSLSAMR 48

Search completed: April 19, 2001, 13:00:00  
 Job time: 118 sec

---

Query Match Score 7; DB 23; Length 46;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 55 PPGLLPR 61  
 Db 40 PPGLLPR 46

RESULT 15  
 US-08-424-550A-493  
 Sequence 493, Application US/08424550A  
 GENERAL INFORMATION:  
 APPLICANT: JOHN N. SIMONS  
 APPLICANT: TAMI J. PILOT-MATIAS  
 APPLICANT: GEORGE J. DAWSON  
 APPLICANT: GEORG G. SCHLAUDER  
 APPLICANT: SURUSH M. DESAI  
 APPLICANT: THOMAS P. LEARY  
 APPLICANT: ANTHONY SCOTT MUERHOF  
 APPLICANT: JAMES C. ERKER  
 APPLICANT: SHERI L. BULJK  
 APPLICANT: ISA K. MUSHAHWAR  
 TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS

